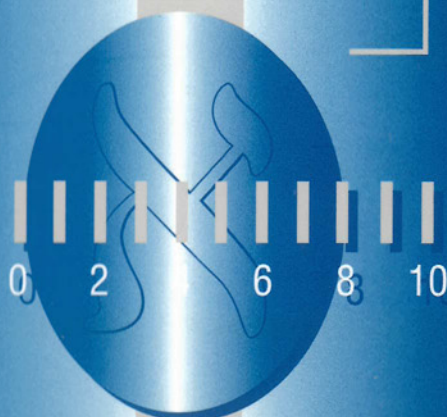


Series on Advances in Mathematics for Applied Sciences – Vol. 66

ADVANCED
**MATHEMATICAL &
COMPUTATIONAL TOOLS
IN METROLOGY VI**

Editors

**P Ciarlini
M G Cox
F Pavese
G B Rossi**



World Scientific

**ADVANCED
MATHEMATICAL &
COMPUTATIONAL TOOLS
IN METROLOGY VI**

Book Editors

Patrizia Ciarlini (Istituto per le Applicazioni del Calcolo, CNR, Roma, Italy),

Maurice G Cox (National Physical Laboratory, Teddington, UK),

Franco Pavese (Istituto di Metrologia "G.Colonnetti", CNR, Torino, Italy),

Giovanni B Rossi (Universita' di Genova, DIMEC, Genova, Italy)

For the first four Volumes see this Series vol. 16 (1994), vol.40 (1996), vol.45 (1997), vol.53 (2000) and vol.57 (2001)



THEMATIC NETWORK "ADVANCED MATHEMATICAL AND COMPUTATIONAL TOOLS IN METROLOGY" (FP5, SOFTTOOLSMETRONET). Coordinator: F Pavese, Istituto di Metrologia "G.Colonnetti" (IMGC) , Torino, IT (*EU Grant G6RT-CT-2001-05061 to IMGC*)

Torino Chairman: F Pavese, Istituto di Metrologia "G.Colonnetti" (IMGC) , Torino, IT

INTERNATIONAL SCIENTIFIC COMMITTEE

George Bonnier, Institut National de Métrologie, Paris, France

Juan Casas-Cubillos, CERN, Geneva

Patrizia Ciarlini, Istituto per le Applicazioni del Calcolo "M.Picone", Roma, Italy

Maurice G. Cox, National Physical Laboratory (NPL-DTI), Teddington, UK

Daniela Ichim, Istituto di Metrologia "G.Colonnetti" (IMGC), Torino, Italy

Rayner Kohler, Bureau International des Poids et Mesures (BIPM), Sèvres

Eduarda Côrte-Real Filipe, Instituto Português da Qualidade (IPQ), Caparica, Portugal

Franco Pavese, Istituto di Metrologia "G.Colonnetti" (IMGC), Torino, Italy

Maria L Rastello, Istituto Elettrotecnico Nazionale "G.Ferraris" (IEN), Torino, Italy

Dieter Richter, Physikalisch-Technische Bundesanstalt (PTB), Berlin, Germany

Gert Rietveld, Nederlands Meetinstituut, Delft, The Netherlands

ORGANISED BY

CNR, Istituto di Metrologia "G.Colonnetti", (IMGC) Torino, Italy

Istituto Elettrotecnico Nazionale "G.Ferraris" (IEN), Torino, Italy

LOCAL COMMITTEE

D.Ferri, D.Ichim, A.Merlone, F.Pavese, I.Peroni, T.Rolandi, F.Sparasci (IMGC),

E.Melli (IEN)

Sponsored by

EU Thematic Network SofTools_MetroNet

IMGC-CNR

IAC-CNR

IEN

Societa' Italiana di Matematica Applicata ed Industriale (SIMAI)

Associazione per lo Sviluppo Scientifico e Tecnologico del Piemonte (ASP), Italy

NPL-DTI, United Kingdom

PTB, Germany

NMI-VSL, The Netherlands

BNM-INM, France

EUROMET

ADVANCED

**MATHEMATICAL &
COMPUTATIONAL TOOLS
IN METROLOGY VI**

Editors

P Ciarlini

CNR, Istituto di Applicazione del Calcolo "M. Picone", Roma, Italy

M G Cox

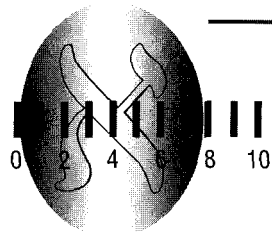
National Physical Laboratory, Teddington, UK

F Pavese

CNR, Istituto di Metrologia "G. Colonnetti", Torino, Italy

G B Rossi

Università di Genova DIMEC, Genova, Italy



World Scientific

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: Suite 202, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

**ADVANCED MATHEMATICAL AND COMPUTATIONAL TOOLS
IN METROLOGY VI**

Series on Advances in Mathematics for Applied Sciences — Vol. 66

Copyright © 2004 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-238-904-0

Foreword

This volume collects the refereed contributions based on the presentation made at the sixth workshop on the theme of advanced mathematical and computational tools in metrology, held at the Istituto di Metrologia “G.Colonnetti” (IMGC), Torino, Italy, in September 2003. The aims of the European Project now supporting the activities in this field were

- To present and promote reliable and effective mathematical and computational tools in metrology.
- To understand better the modelling, statistical and computational requirements in metrology.
- To provide a forum for metrologists, mathematicians and software engineers that will encourage a more effective synthesis of skills, capabilities and resources.
- To promote collaboration in the context of EU Programmes, EUROMET and EA Projects, MRA requirements.
- To support young researchers in metrology and related fields.
- To address industrial requirements.

The themes in this volume reflect the importance of the mathematical, statistical and numerical tools and techniques in metrology and also keeping the challenge promoted by the Meter Convention, to access a mutual recognition for the measurement standards.

Torino, February 2004

The Editors

This page intentionally left blank

Contents

<i>Foreword</i>	v
Estimation of Precision and Uncertainty of a Calibration Artefact for CMMs <i>S. D. Antunes and M. A. F. Vicente</i>	1
Uncertainty in Semi-Qualitative Testing <i>W. Bremser and W. Hässelbarth</i>	16
Processing the Coherent Anomalies on Digitalized Surfaces in Wavelet Domain <i>P. Ciarlini and M. L. Lo Cascio</i>	24
Least Squares Adjustment in the Presence of Discrepant Data <i>M. G. Cox, A. B. Forbes, J. L. Flowers and P. M. Harris</i>	37
Harmonization of Correlated Calibration Curves with an Application to the Analysis of Natural Gases <i>M. G. Cox, S. Kamvissis, M. J. T. Milton and G. Vargha</i>	52
Parametrized Approximation Estimators for Mixed Noise Distributions <i>D. P. Jenkinson, J. C. Mason, A. Crampton, M. G. Cox, A. B. Forbes and R. Boudjemaa</i>	67
Algorithms for the Calibration of Laser-Plane Sensors on CMMs <i>C. Lartigue, P. Bourdet, L. Mathieu and C. Mehdi-Souzani</i>	82
Some Differences between the Applied Statistical Approach for Measurement Uncertainty Theory and the Traditional Approach in Metrology and Testing <i>C. Perruchet</i>	98
Metrology Software for the Expression of Measurement Results by Direct Calculation of Probability Distributions <i>G. B. Rossi, F. Crenna and M. Codda</i>	106

Feasibility Study of Using Bootstrap to Compute the Uncertainty Contribution from Few Repeated Measurements <i>B. R. L. Siebert and P. Ciarlini</i>	122
Recursive and Parallel Algorithms for Approximating Surface Data on a Family of Lines or Curves <i>G. Allasia</i>	137
Process Measurement Impact on the Verification Uncertainty <i>J. Bachmann, J. M. Linares, S. Aranda and J. M. Sprauel</i>	149
On the In-Use Uncertainty of an Instrument <i>W. Bich and F. Pennecchi</i>	159
Automatic Differentiation and Its Application in Metrology <i>R. Boudjemaa, M. G. Cox, A. B. Forbes and P. M. Harris</i>	170
Usage of Non-Central Probability Distributions for Data Analysis in Metrology <i>A. Chunovkina</i>	180
Implementation of a General Least Squares Method in Mass Measurements <i>J. Hald and L. Nielsen</i>	189
The GUM Tree Design Pattern for Uncertainty Software <i>B. D. Hall</i>	199
Statistical Hypotheses Testing for Phase Transition Identification in Cryogenic Thermometry <i>D. Ichim, I. Peroni and F. Sparasci</i>	209
The Impact of Entropy Optimization Principles on the Probability Assignment to the Measurement Uncertainty <i>G. Iuculano, A. Zanolini and G. Pellegrini</i>	219
Stochastic Processes for Modelling and Evaluating Atomic Clock Behaviour <i>G. Panfilò, P. Tavella and C. Zucca</i>	229
Compound-Modelling of Metrological Data Series <i>F. Pavese</i>	240

Homotopic Solution of EW-TLS Problems <i>M. L. Rastello and A. Premoli</i>	251
Pooled Data Distributions: Graphical and Statistical Tools for Examining Comparison Reference Values <i>A. G. Steele, K. D. Hill and R. J. Douglas</i>	264
Numerical Uncertainty Evaluation for Complex-Valued Quantities: A Case Example <i>L. Callegaro, F. Pennecchi and W. Bich</i>	274
Bayesian Approach to Quantum State Tomography <i>S. Castelletto, I. P. Degiovanni, M. L. Rastello and I. Ruo Berchera</i>	279
Simulation of Charge Transfer in a Tunnel Junction: Approaching the Sub- <i>e</i> Scale <i>G. E. D'Errico</i>	282
Validation of Calibration Methods — A Practical Approach <i>E. Filipe</i>	285
Comparison of LS Techniques for the Linear Approximation of Data Affected by Heteroschedastic Errors in Both Variables with Uncertainty Estimation <i>D. Ichim, P. Ciarlini, F. Pavese, A. Premoli and M. L. Rastello</i>	293
Noise Correction for Surface Measurements <i>H. Haitjema and M. A. A. Morel</i>	299
Evaluation of Uncertainty of Standard Platinum Resistance Thermometer at National Laboratory Level <i>M. J. Korczynski, A. Hetman, A. Szmyrka-Grzebyk and P. Fotowicz</i>	302
A New Approach to the Presentation of the Result Measurements in Virtual Instruments <i>M. J. Korczynski and A. Hetman</i>	309
A Hybrid Method for ℓ_1 Approximation <i>D. Lei and J. C. Mason</i>	314

Interpolation Equations for Industrial Platinum Resistance Thermometers <i>P. Marcarino, P. P. M. Steur and A. Merlone</i>	318
From the Fixed Points Calibration to the Certificate: A Completely Automated Temperature Laboratory <i>A. Merlone, P. Marcarino, P. P. M. Steur and R. Dematteis</i>	323
A New Off-Line Gain Stabilisation Method Applied to Alpha-Particle Spectrometry <i>S. Pommé and G. Sibbens</i>	327
Development of Software for ANOVA that Can Generate Expressions of Variance Expectations <i>H. Tanaka, T. Kamoshita and K. Ehara</i>	330
Template Matching in Passive Sonar Recognition <i>J. L. Terry, D. A. Turner and J. C. Mason</i>	333
Fast Computational Alternative to Monte Carlo Simulation of an Output Distribution <i>R. D. Willink and B. D. Hall</i>	337
<i>Short Course on Uncertainty Evaluation</i> <i>M. G. Cox</i>	342
<i>Software Requirements in Legal Metrology: Short Course Held Adjacent to the Conference</i> <i>D. Richter</i>	345
<i>Author Index</i>	349

ESTIMATION OF PRECISION AND UNCERTAINTY OF A CALIBRATION ARTEFACT FOR CMMS

S. D. ANTUNES

Instituto Superior Técnico, Av^a Rovisco Pais, 1049-001 Lisboa, Portugal

E-mail: santunes@ist.utl.pt

M. A. F. VICENTE

Universidade de Coimbra, Departamento de Matemática da FCTUC, Apartado 3008

3001-454 Coimbra, Portugal

E-mail: vicente@mat.uc.pt

A new kind of artefact, based on a modification of the hexapod machine's well-known structure, has been introduced by Antunes, S. D. *et al* in [1], in order to determine the global errors of coordinate measuring machines. Here we are presenting results from validation of the technique: using a self-calibrated method and modeling the reference value for calibration based on laser trilateration.

1. Introduction

Uncertainty, as defined in the ISO Guide to the Expression of Uncertainty in Measurement (GUM) [4] and in the International Vocabulary of Basic and General Terms in Metrology (VIM), is a parameter, associated with the result of a measurement, which characterizes the dispersion of the values that could reasonably be attributed to the measurand.

Calibration and numerical error correction of coordinate measuring machines (CMMs) require an error behavior mathematical model and methods to assess the errors. Traceability of CMMs must be based on traceability chains and commonly accepted methods for uncertainty evaluation [3].

2. Calibration Artefacts for CMMs

Different artefacts (geometrical gauges) can be used to perform the calibration of large CMMs (with one or more axis length bigger than 2 m). Between the calibrated artefacts, we can refer the 2D lightweight ball plates (with carbon fiber rod structure and ceramic spheres), the 2D disassemblable ball plates (with L-shape, for example, made of carbon fiber tubes) or the 1D disassemblable multi-ball bars (of carbon fiber pipes), as used by [2]. Alternatively, to perform this task, the calibration can be made with the use of an arrangement of laser

interferometers and level meters or straight edges to access the full error analysis, but these methods require expensive tools, are time consuming and need especially skilled personnel. In addition, it is possible to use uncalibrated artefacts, but it is necessary to have absolute length standards to determine scale factors. In this last case, it is necessary to place the artefact in substantially more positions than when using calibrated artefacts.

For small and medium size coordinate measuring machines there are several artefact-based methods for full error analysis, which is a prerequisite for establishing traceability. Like the case of large CMMs, there are calibrated and uncalibrated artefacts for this purpose. The most common calibrated artefacts are the gage blocks, step gages, the ball bars, the ball plates and the hole plates^a. In addition, laser interferometers are also used in order to perform the measurement of linear accuracy, straightness and angularity checks. Additionally, there are uncalibrated devices, like the artefact presented in [1], which is the object of the main study described in this text.

Artefacts are effective for local and relative calibrations of CMMs. We have compared four different types of reference artefacts [5]:

- the single sphere (see figure 1);
- the step gauge block (see figure 2);
- the ball plate (see figure 3);
- the hole bar (see figure 4).

The single sphere is used to calibrate the probe of the CMM and a good metrological sphere of known and calibrated diameter is needed. All the other listed artefacts are used to calibrate the CMM, in order to find systematic errors on the CMM's measurement volume.

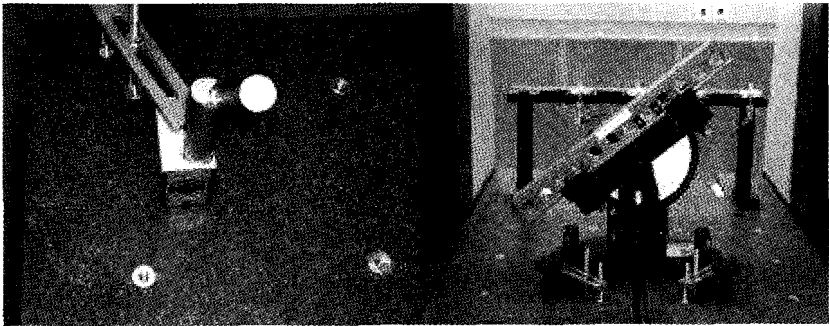


Figure 1. The sphere.

Figure 2. The step gauge block.

^a Like the presented artefact in <http://www.lgg.com/html/calibcmm.html>.

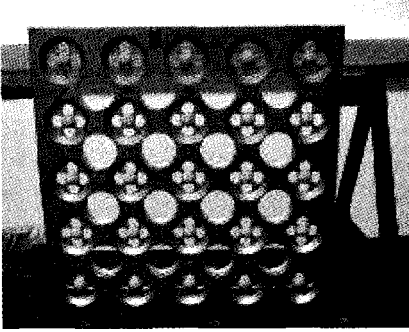


Figure 3. The ball plate.

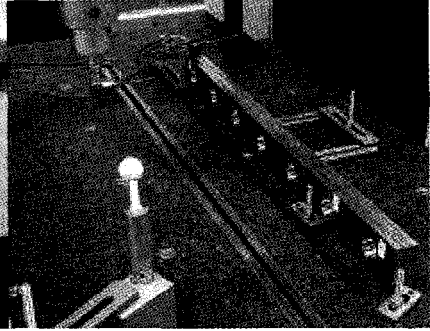


Figure 4. The hole bar.

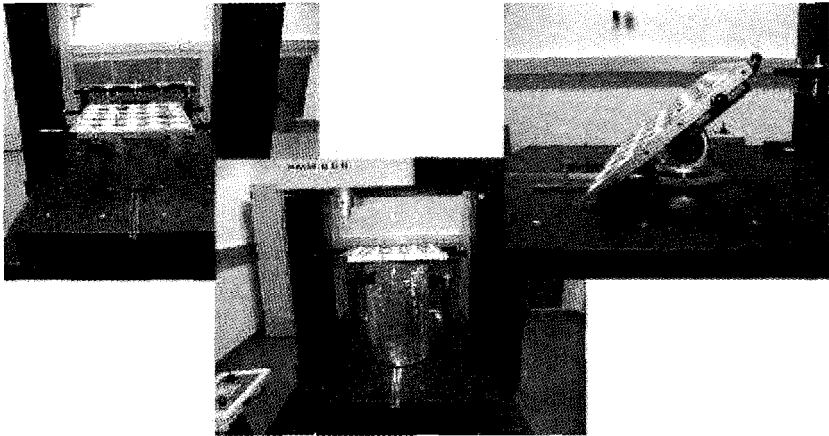


Figure 5. Trying to calibrate CMM's volume measurement with different supports for the ball plate.

The step gauge block enables a linear calibration in a direction perpendicular to the reference surface of the gauges.

The ball plate allows the calibration in the plane of the center of the balls constituting the ball plate (the reference points are the center of the balls in its plane).

The hole bar allows a linear calibration in the plane of one surface of the holes (the reference points are the center of the holes in that plane). But the hole bar also enables the measurement of local errors measuring the position of the axis of reference holes in the bar.

We tried to solve the problem of volume calibration with a ball plate, developing specific supports for using the ball plate (see figure 5), but we always had the problem on how correlate measurements with ball plate calibration.

Those artefacts are effective for local and relative calibrations, but are not appropriate for measuring all calibrating parameters of a CMM. In order to solve those problems, a new kind of artefact, based on a modification of the known structure of the hexapod machine, has been introduced in [1]. The proposed artefact was similar to a tripod system constituted by three links connecting three fixed spheres to a mobile sphere, with the mobile sphere acting as the reference point for CMM calibration (see figure 6, from [1]).

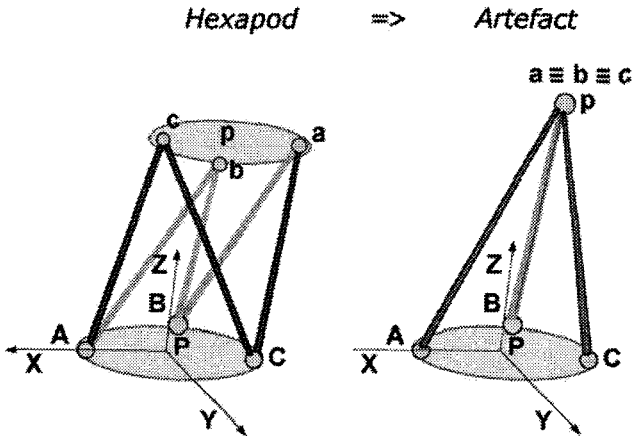


Figure 6. The calibrating artefact, hexapod based.

This kind of artefact needs to be validated, for the purpose of verifying its adequacy. Its validation and the measurement techniques use data fusion applied to non-linear and non-observable systems.

Comparing the calculated position of the moving sphere center with the corresponding measured position by the CMM in the measurement volume does the global calibration of CMMs. Only length variations are used to obtain the calculated position and these length variations are measured by three miniature laser interferometers installed inside the telescopic links (see figures 7 and 8). The artefact uses a self-calibrated method and its modelling is based on laser trilateration.

The problem of global calibration can be divided into two steps:

1st step - *prediction* (geometry identification): knowledge about system geometry (localization of fixed points) and tracking (following mobile point) are used to obtain geometry identification (that is done by finding the initial length of the links l_i , the coordinates of the mobile point x_j, y_j , and the coordinates of the fixed points $x0_j, y0_j$, that minimize:

$$\sum_{i=1}^m \sum_{j=1}^n \left[(l_i + dl_{ij}) - \sqrt{(x0_i - y_j)^2 + (y0_i - y_j)^2} \right]^2 \quad (1)$$

with dl_{ij}^{β} representing laser interferometer measurements);

2nd step - *update* (system identification): after tracking and identifying geometry, the problem is the identification of the mobile point position, in real time, using 3D laser trilateration.

3. Modeling Artefact for Global Calibration of CMMs

Let us consider three fixed spheres P_1 , P_2 and P_3 , each one connected to a mobile sphere Q by a telescopic link (see figure 8). If we know the lengths of the lines connecting the center of the fixed spheres to the mobile sphere perturbed by known noise and by random noise:

$$L_i = L_{0i} + p_{ci} + p_{ui}, \quad (2)$$

from trilateration, we can find the coordinates of the mobile sphere $Q(X, Y, Z)$ and also the uncertainty of each coordinate (u_X, u_Y, u_Z) .

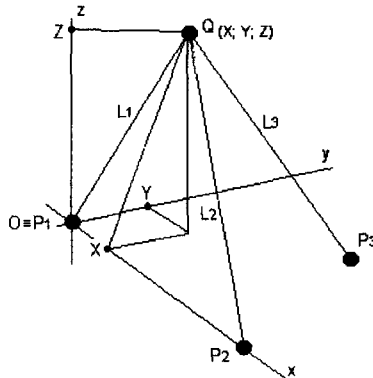


Figure 7. Reference coordinates $Oxyz$.

Considering the coordinates from the fixed spheres centers $P_1(X_1, Y_1, Z_1)$, $P_2(X_2, Y_2, Z_2)$, $P_3(X_3, Y_3, Z_3)$ and assuming $Z_1 = Z_2 = Z_3 = 0$, we obtain:

^b The measurements of links increments, dl_{ij} , are perturbed, at least, by ambient humidity, ambient temperature and atmospheric pressure. Also, measured dl_{ij} are estimated with an uncertainty composed by the uncertainties of measured perturbations, which probability density functions are of unknown type gaussian/non-gaussian and, certainly, constituting processes that may be stationary/non-stationary.

$$X = \frac{X_1^2 - X_2^2 + L_2^2 - L_1^2}{2 \cdot (X_1 - X_2)},$$

$$Y = \frac{Y_1 + Y_3}{2} \cdot \frac{X_1^2 - X_3^2 + L_3^2 - L_1^2}{2 \cdot (Y_1 - Y_2)} \cdot \frac{(L_2^2 - L_1^2 + X_1^2 - X_2^2) \cdot (X_3 - X_1)}{2 \cdot (X_1 - X_2) \cdot (Y_1 - Y_3)} \quad (3)$$

and

$$Z = \sqrt{L_1^2 - (X - X_1)^2 - (Y - Y_1)^2}.$$

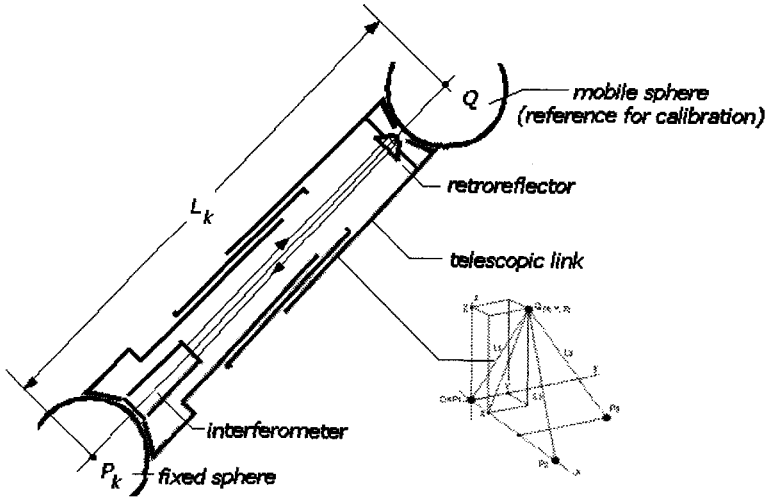


Figure 8. Telescopic links.

Or, particularly, if we consider the specific coordinates from figure 7, $P_1(0,0,0)$, $P_2(X_2,0,0)$ and $P_3(X_3,Y_3,0)$, we get:

$$X = \frac{L_1^2 - L_2^2 + X_2^2}{2 \cdot X_2},$$

$$Y = \frac{X_3^2 + Y_3^2 + L_1^2 - L_3^2 - X_3 \cdot \frac{L_1^2 - L_2^2 + X_2^2}{X_2}}{2 \cdot Y_3} \quad (4)$$

and

$$Z = \sqrt{L_1^2 - X^2 - Y^2}.$$

We use those equations in order to find reference values for the artefact and the first set of equations to find, following the GUM, the uncertainty for the reference coordinates X , Y and Z knowing the uncertainties for the center of the

fixed spheres $X_1, Y_1, Z_1, X_2, Y_2, Z_2, X_3, Y_3$ and Z_3 , and also for the length of the links L_1, L_2 and L_3 .

Following the GUM, the uncertainty u_Y for a dependent measure $Y = Y(X_1, X_2, X_3, \dots, X_n)$, a function of n uncorrelated measurements X_i with uncertainties given by u_{X_i} , is defined by:

$$u_Y = \sqrt{\sum_{i=1}^n \left(\frac{\partial Y}{\partial X_i} \cdot u_{X_i} \right)^2}. \quad (5)$$

In order to bypass the derivatives $\frac{\partial Y}{\partial X_i}$, we are using the difference resulting from increments on X_i equal to the uncertainty u_{X_i} . Consequently, u_Y is given by:

$$u_Y = \sqrt{\sum_{k=1}^n [Y(X_1, X_2, \dots, X_k, \dots, X_n) - Y(X_1, X_2, \dots, X_k + u_{X_k}, \dots, X_n)]^2}. \quad (6)$$

Two problems exist relating to the length of each link L_i and also to the coordinates for the fixed spheres $P_1(X_1, Y_1, Z_1)$, $P_2(X_2, Y_2, Z_2)$ and $P_3(X_3, Y_3, Z_3)$. With the laser interferometers we only are measuring length variation. We need to estimate the real length for the links, and also the coordinates for fixed spheres. We assume that $X_1 = Y_1 = Z_1 = Y_2 = Z_2 = Z_3 = 0$, using the reference coordinates from figure 7. All the other values, and also the uncertainties for all values, are calculated by self-calibration.

4. Self-calibration of the Artefact

Self-calibration (system identification) is done by moving the mobile sphere over the plane of the fixed spheres and minimizing the cost function c :

$$c = \sum_{k=0}^m (a_{1_k} + a_{2_k} + a_{3_k}) \quad (7)$$

where $a_{1_k}, a_{2_k}, a_{3_k}$ are given by the set of $3(m+1)$ equations, for $k=0, \dots, m$:

$$\begin{aligned} a_{1_k} &= (X_{k_est} - X_{1_est})^2 + (Y_{k_est} - Y_{1_est})^2 + (Z_{k_est} - Z_{1_est})^2 - (L_{1_est} + d_{1_k})^2 \\ a_{2_k} &= (X_{k_est} - X_{2_est})^2 + (Y_{k_est} - Y_{2_est})^2 + (Z_{k_est} - Z_{2_est})^2 - (L_{2_est} + d_{2_k})^2 \\ a_{3_k} &= (X_{k_est} - X_{3_est})^2 + (Y_{k_est} - Y_{3_est})^2 + (Z_{k_est} - Z_{3_est})^2 - (L_{3_est} + d_{3_k})^2 \end{aligned} \quad (8)$$

with $X_{k_est}, Y_{k_est}, Z_{k_est}$ the estimated coordinates of the mobile sphere in the k th position, $L_{1_est}, L_{2_est}, L_{3_est}$ the estimated lengths for the three links at the initial position for $k=0$, and $d_{1_k}, d_{2_k}, d_{3_k}$ the measured length variations of the three links.

For $m+1$ positions of the mobile sphere, we get $3(m+1)$ equations with equal number of measurements and with $3m+13$ unknowns ($3m+3$ coordinates

for the mobile sphere positions, 9 coordinates for the fixed spheres position and 3 initial lengths for the links).

The problem has many solutions, and the fixed sphere's location enables a quick evaluation of the validity of the solution. In figure 9 we present four possible solutions for the same inputs, resulting from the evident axis symmetrical solutions. The optimal solution search is done using the optimization toolbox of *Matlab*: starting with a constrained method in order to fix the solution in the adequate region and concluding with an unconstrained method to obtain a quicker and better solution.

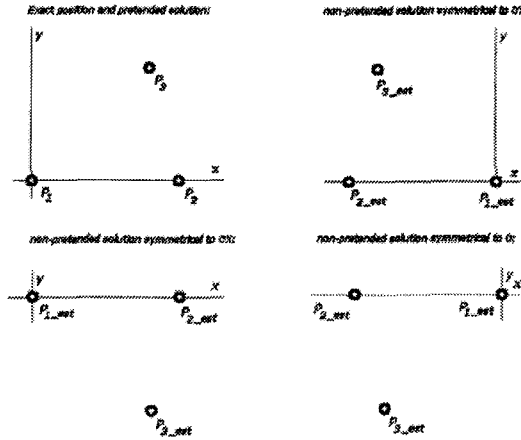


Figure 9. Those figures present possible evident solutions for self-calibration.

The cost function value is used to obtain the estimated uncertainty for the estimated parameters.

5. Simulating Calibration Artefact

In order to verify the adequacy of use of the developed device, some numerical simulations were made, using the software *Matlab* (version 6.5 R13). The programs are divided into two major parts: in the first one, the purpose is to realize the self-calibration of the system and, in the second, using the results for the adjusted parameters obtained in the first part, using trilateration, estimated coordinates and the correspondent uncertainty are obtained for some reference points.

The program needs initial parameters that are the minimum and the maximum link length, the link's length variation noise amplitude, the number of

constrained steps (necessary to constrain the area where the solutions could be found, using the optimization method) and forbidden area radius (where no points can be placed, corresponding to a physical limitation of the artefact). Table 1 summarizes the parameters (in relative units) introduced in the presented examples.

Table 1. Parameters used for simulations.

Parameters		Simulation		
		1 st	2 nd	3 rd
	maximum link length	20	20	20
	links length noise amplitude	0	0.01	0.30
	constrained steps	5	5	5
	forbidden area radius	0.6	0.6	0.6

After the introduction of the initial parameters, the user chooses, in a graphical way using the mouse pointer, the three fixed spheres' positions and the various moving sphere positions. In figure 10 these positions are presented for the 2nd simulation, similar to the other two simulations realized (note that in the space of measurement there are some regions of shadow, corresponding to zones that are unreachable by the moving sphere). As the user introduces a new position, the parameters are recalculated using all information available.

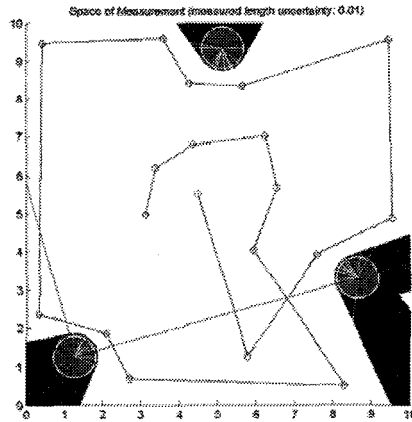


Figure 10. Example of fixed spheres position and mobile sphere trajectory during self-calibration.

After introducing all moving sphere positions, the first part of the program is finished and the results of self-calibration are presented, in a graphical and in a numerical form. In table 2 we present the final numerical results of

self-calibration for the 2nd simulation corresponding to links length noise amplitude equal to 0.01.

The results, in a graphical form, are presented in figure 11, which shows the real positions *versus* estimated positions of the fixed points, the variations of the link lengths, the evolution of the cost function that is minimized during the program and the evolution of the estimated values for the initial link lengths.

For the 1st and the 3rd simulations the graphical results of self-calibration are shown in figures 12 and 13.

Table 2. Results of self-calibration with links length noise amplitude 0.01.

	reference value	estimated value	deviation
initial links lengths	4.14524605719157	4.14762057356121	0.00237451636964
	5.77216978117969	5.76981912447491	-0.00235065670478
	4.80120164387303	4.81250457805668	0.01130293418365
x of 2 nd fixed point	7.64741472782958	7.64545798330029	-0.00195674452929
x of 3 rd fixed point	5.87877399052966	5.89051558442518	0.01174159389553
y of 3 rd fixed point	6.74274716485670	6.75540982406257	0.01266265920588
1 st moving point x	2.76878359671112	2.77157430333704	0.00279070662592
1 st moving point y	3.08494769311341	3.08807624576027	0.00312855264686
2 nd moving point x	3.35102638810543	3.36071345716409	0.00968706905866
2 nd moving point y	4.19778270569769	4.19928162592284	0.00149892022515
...

After the self-calibration, it is possible to introduce in the program some values of 3D reference points coordinates and compare them with the correspondent estimated coordinates, obtained from the final parameters (given by the self-calibration process) using trilateration. In addition, those results have a corresponding estimated uncertainty, a function of the uncertainty of the links' lengths in table 3 are summarized some results for 4 reference points obtained in the three simulations realized.

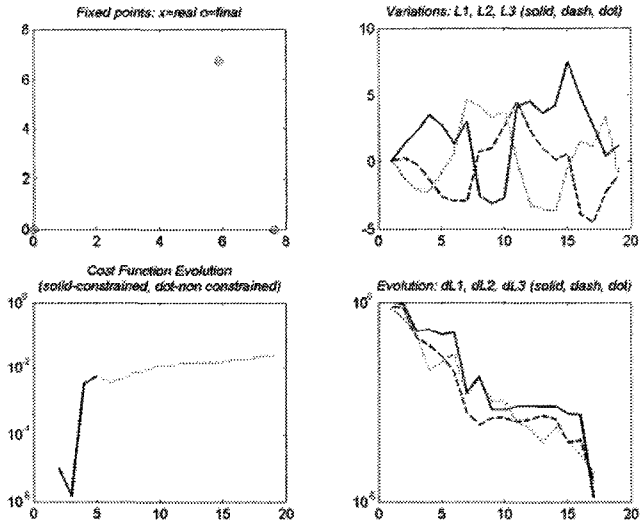


Figure 11. Self-calibration with link's length noise amplitude = 0.01.

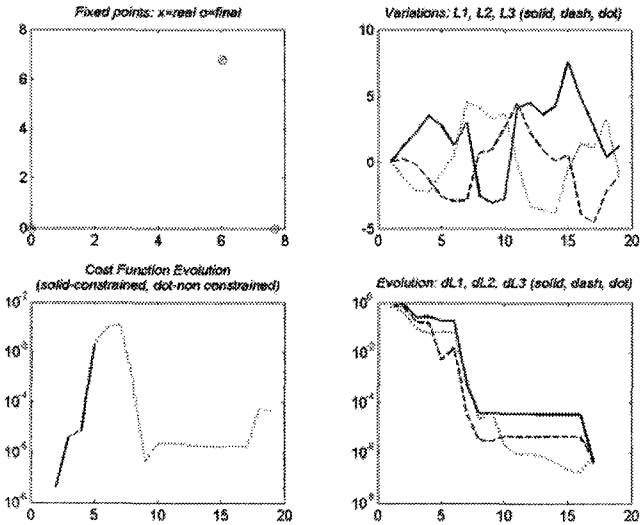


Figure 12. Self-calibration with link's length noise amplitude = 0.0.

6. Reference Values for CMM Calibration – Uncertainty Propagation

Coordinates, and uncertainties for those coordinates, are the reference values for CMMs calibration. The coordinates and the respective uncertainties are calculated in real time, by a method based on laser interferometer measurement of the links length variation.

For that purpose, we tried several Data Fusion solutions such as the *Kalman* Filter, the Extended *Kalman* Filter, the covariance intersection, etc. In this paper, we present the results obtained using the standard definition for uncertainty propagation (due to its extension and unsolved small issues, further results will be presented later).

In figure 14, we represent some of the estimated uncertainty values for X , Y and Z coordinates of the moving sphere in some regions of the measurement space, for a link length noise amplitude equal to 0.01. As we can see, the final uncertainty in each coordinate, not only depends on the link length uncertainty value, but also depends on the value of that coordinate and on the values of the other two coordinates.

From figure 14, we can also see that the uncertainties of X and Y slowly increase when the coordinate Z from the mobile sphere increases. The uncertainty for the Z coordinates increases during trilateration for small values of Z and reduces when Z is increasing. Those values are not acceptable and trilateration only may be used for values of Z not equal to 0. When the mobile sphere is over the plane of the fixed spheres, only the X and Y coordinates must be calibrated, as in the case of using the ball plate artefact.

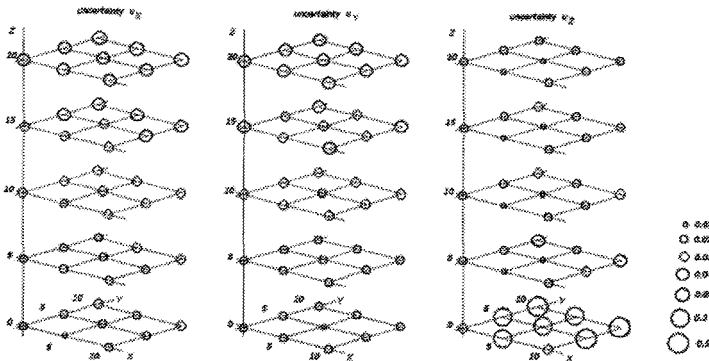


Figure 14. Moving sphere coordinates uncertainty in several regions of the measurement space.

7. Conclusions

The measurements made with the use of the described artefact have errors and a previously unknown uncertainty. In this text, above, we presented a technique to incorporate the measured values uncertainties in order to obtain the calibration results' final uncertainties (using data fusion as a learning system, meaning that the optimal uncertainty estimation is constantly adapted during self-calibration of the device). However, this is not the final answer for this problem of artefact validation: we are developing and comparing different techniques in non-linear and non-observable state estimation.

The results obtained for the final calibration uncertainty using those techniques lead us to the following conclusions. First, in order to decide among the various techniques, we need to take into account the amount of computer memory required for identification, the accuracy of the method, the difficulties in implementation and also the ability to represent multimodal probability density functions (PDFs).

In table 4, where *KF* represents the *Kalman* filter, *EKF* represents the extended *Kalman* filter, *CI* represents the covariance intersection, *CP* represents the covariance propagation, *Markov* represents the interval functions with fixed cells, probability grids or maximum entropy estimation and, finally, *PF* represents various particle filters (the set constituted by *Monte Carlo* localization, *bootstrap*, condensation algorithm and fittest algorithm), we summarize the comparison of the different techniques studied.

Table 4. Comparing calculating data fusion techniques for uncertainty.

	<i>KF</i>	<i>EKF</i>	<i>CI</i>	<i>CP</i>	<i>Markov</i>	<i>PF</i>
Starting geometry knowledge	good estimation		from scratch			
Time of response	more	more	more	more	more	less
Memory consuming	less	less	less	less	more	less
Ability to represent multimodal PDFs	less	less	more	more	more	more

As a conclusion, we can state that the particle filters seem to be the best methods for our goal. However, the ability to represent multimodal PDFs seems to be very restrictive and a better analysis need to be realized taking into

account the costs involved (time response, memory consuming, etc.). In other way, the *Kalman* filter and the extended *Kalman* filter need a starting geometry knowledge that can be estimated, measuring the position of fixed points and changing coordinates following the artefact coordinate system.

The results given by using the developed *Matlab* routines based on the above make us conclude that the self-calibration of the artefact works well. The modelling of the artefact allows us to develop the real artefact that we are starting to build and, in the near future, we will have results from that artefact.

References

1. S. D. Antunes, P. M. Antunes and M. A. F. Vicente; *Local and Global Calibration of Coordinate Measuring Machines*; in *Advanced Mathematical and Computational Tools in Metrology V*; Series on Advances in Mathematics for Applied Sciences Vol.57, World Scientific (2001) pp 16-22.
2. L. Arriba, E. Trapet, M. Bartscher, M. Franke, A. Balsamo, G. Costelli, S. Torre, F. Kitzsteiner and F. San Martín; *Methods and Artifacts to Calibrate Large CMMs*, in Proceedings of the 1st international EUSPEN conference, Bremen, Germany (1999) pp 391-394.
3. M. Bartscher, K. Busch, H. Schwenke, E. Trapet and F. Wäldele; *Artefact based Traceability Concept for Large Co-ordinate Measuring Machines (CMMS)*, in Proceedings of the 9th METROLOGIE Conference, Bordeaux, France (1999) pp 158-161
4. ISO; *Guide to the Expression of Uncertainty in Measurements (GUM)*, Geneva, Switzerland (1995).
5. J. F. Silva; *Artefactos para a calibração de máquinas de medir por coordenadas*, rapport of *Projecto de Produção*, under the supervision of S. D. Antunes, IST-DEM-STM (2001).

UNCERTAINTY IN SEMI-QUALITATIVE TESTING

W. BREMSER AND W. HÄSSELBARTH

Federal Institute for Materials Research and Testing (BAM)

Unter den Eichen 87, 12205 Berlin, Germany

E-mail: wolfram.bremser@bam.de

The problem of assigning uncertainty-like statements to *qualitative* test results still remains widely unsolved, although the volume of not purely quantitative testing and analysis, and their economic impact are immense. A pragmatic approach is developed for the assessment of measurement uncertainty for procedures where a complex characteristic (e.g. of a material) discontinuously changes in dependence on one or more recordable, continuous quantitative parameters, and the change of the characteristic is assessed by judgement (which may be instrument-assisted). The principles of the approach are discussed, and an application example is given.

1. Introduction

While calculation, assessment, and statement of measurement uncertainty in *quantitative* testing is required by international standards [1], supervised by accreditation bodies [2], and guided by basic documents [e.g. 3] such that the process of implementation of the measurement uncertainty concept has been put on the wheels in calibration and testing laboratories throughout the world, the problem of assigning uncertainty-like statements to *qualitative* test results remains widely unsolved. ILAC states that for qualitative testing, consideration is still being given as to how uncertainty of measurement applies in such cases, and neither the GUM [3] nor the EURACHEM/CITAC Guide [4] provide any guidance on how to tackle the problem.

On the other hand, the volume of not purely quantitative testing and analysis is immense, and so is the economic impact of decisions taken on the basis of qualitative results. Therefore, it seems reasonable to develop feasible approaches. Fortunately, nowadays a very large group of qualitative testing procedures are not as qualitative as it may seem at a first glance. Rare are the situations where testing procedures are purely descriptive, or solely depend on human senses. Most frequently, one will have a situation where a complex characteristic (e.g. of a material) discontinuously changes in dependence on one or more recordable, continuous quantitative parameters. Examples are e.g. material breakdown or detonation of explosive mixtures.

Whether or not the characteristic has changed at a certain set of parameter values is assessed by judgement (e.g. pattern recognition, colour change) which may be instrument-assisted. Such procedures are semi-qualitative in the sense that the judgement is qualitative, but the result of the test is nevertheless a fully quantified (set of) parameter(s). A certain fuzzyness of judgement induces an uncertainty in quantification.

2. Principle of the approach

Very much in general, basic approaches like method repeatability and method reproducibility remain also valid for semi-qualitative test procedures. Any testing laboratory may create an estimate of its internal method scatter simply by assessing the results of an appropriate number of repetitions. Such experiments can be carried out under either repeatability (controlled, constant environmental and other influential parameters, same operator) or reproducibility conditions (maximum variation of all influential parameters within the limits of method specification, and different operators).

Laboratory intercomparisons may be designed and executed in the very same way. Average lab precision and average between-lab bias may be calculated from the intercomparison results using established algorithms.

Beyond the limits of these more classical approaches, a "component-by-component" approach is also feasible which will be even closer to the spirit of the GUM. The basic element of this approach is also repetition. The characteristic to be tested is modelled as a Bernoulli random variable depending on a set of fully quantified parameters

$$\zeta = f(k_1, k_2, \dots, k_n). \quad (1)$$

The variable may take only two values: 0 (fail) or 1 (success). During the test, the status of ζ in dependence on \mathbf{k} is assessed by judgement, and k_c determined as the value at which the status of ζ changes. Normally, there will be a region of certainty where the outcome of the experiment is undoubtedly 0 or 1, and a region in the vicinity of the threshold k_c where the outcome is uncertain, mainly due to the fuzzyness of judgement.

Within this region, replicate measurements at each set of parameter values are taken, and the parameter values are increased (or decreased) in a stepwise manner according to the actual test requirements. With n replicates, ζ at each sampling point should follow a $B(n, p)$ binomial distribution with known mean and variance. The (changing) probability parameter p is estimated as the frequency of occurrence (or non-occurrence) of the event as observed, and the corresponding variances $\text{var}(p)$ are calculated accordingly.

A unit step function (with the step at k_c) is fitted to the frequency-of-occurrence data. The variable parameter for the fit is k_c . Different minimisation criteria are feasible, namely

i) the normal SSD = \min criterion known from OLS regressions. Due to the discrete spacing of the sampling and the infinitive first derivative of the Heaviside function at k_c , this criterion provides multiple solutions in an interval between two sampling points.

ii) the criterion of equilibrated (i.e. equal) cumulated probabilities (quantiles) on both sides of the step. This criterion provides a unique solution for k_c , which will fall into the interval according to i) and, thus, also satisfies the minimum residual SSD requirement of an OLS fit.

The criterion according to ii) should be preferred. Since at both sides of the step estimates of an average occurrence probability are only available at the discrete sampling points, quantiles may only be calculated with an additional assumption on an appropriate envelope function. Detailed information on the envelope is rather unavailable, but pragmatic approaches can be used including simple frequency polygons which connect the measured data points in a segment-wise manner by straight lines. A disadvantage of this pragmatism is that the areas beneath and above the frequency polygon will normally be different (i.e. the strict $p = 1 - q$ rule is violated), but since one will use in practice only one type of the areas (the cumulated p) on both sides of the step, the method will provide a sensible estimate for the equality point.

It may also seem reasonable to fit the quantile of the normal distribution to the data obtained for p . On the one-hand side, one aims at finding estimates for the variability (scatter) of one (or a couple of) value(s) k_c , which is/are influenced by a considerable number of influential factors such that an overall distribution close to the normal distribution can be assumed. On the other hand, this coincides with the philosophy of the GUM [3], which also recommends the transformation of different, most often only assumed distributions to Gauss distributions, and the corresponding measures of scatter to standard uncertainties.

For the given set of data points $[p_i, k_i]$ one has to solve the minimisation problem

$$\sum [p_i - \text{quant}(k_i, k_c, s)]^2 = \min \quad (2)$$

(with quant being the quantile of the normal distribution) with respect to the parameters k_c and s . This automatically generates the standard uncertainty, the confidence interval is calculated from its multiplication by the corresponding

(one-sided) t factor. The above described approach may be further refined into a weighted (in y) regression using the inverse variances of the p_i as weights.

With or without such weighing in y , one is still not completely done since in a last step, the quantifiable uncertainties of the k_i should be taken into account by uncertainty propagation [5]. Therefore each k_i is separately altered by $\pm \frac{1}{2}u(k_i)$, and the minimisation problem solved for k_c with a fixed s . The contribution of the uncertainty in k_i to the total uncertainty is obtained as

$$u(k_c; k_i) = k_c(k_i + \frac{1}{2}u(k_i)) - k_c(k_i - \frac{1}{2}u(k_i)) \quad (3)$$

and will accordingly be summed up for all sampling points. The above described approach will now be exemplified.

3. Example: Explosion limit determination

The determination of explosion limits according to prEN 1839 is a semi-qualitative testing procedure in the sense described here. The characteristic under investigation is the detonation of an potentially explosive gas mixture. It is assessed by a pattern recognition-like procedure based on the decision whether a flame separation or the formation of a tall aureole took place following ignition. The result is a pass/fail decision. Fully quantifiable influential variables k are

i) the gas composition (flammable gas in air) expressed as a fraction or concentration (in the following: mol%).

ii) the temperature. It is assumed that the characteristic under investigation functionally depends on the temperature. However, for the aims of this study one may restrict all considerations to one (e.g. a reference) temperature T_0 . Obviously, this temperature can also be measured and controlled only with a non-zero uncertainty, but provided the functional relationship is known (e.g. from systematic investigations) this uncertainty may easily be propagated to the final result.

iii) the temperature gradient in the detonation vessel. Without any further information on its influence on, or a possible insignificance (e.g. within certain specification limits) for, the final experimental result this influential factor cannot be taken into account properly.

iv) the ignition energy. There are good reasons for the assumption that the dependence of detonation occurrence on this factor is of a threshold type. Pro-

vided the ignition energy is held well above this threshold, its influence on the uncertainty of the experimental result may be neglected.

Taking the above into account, the problem is reduced to a functional relationship of the binomial event occurrence on the gas composition x (mole fraction).

This composition was altered for a certain gas mixture within the range between 13 and 14.5 mol%, and 10 replicate observations were carried out at each sampling point. Each observation was assessed either as failure or success, and results as displayed in table 1 were obtained.

Table 1. Observed number of ignitions at different gas compositions

#	x mol%	failure no ignition	success ignition	p
1	13.1	10	0	0
2	13.4	9	1	0.1
3	13.6	7	3	0.3
4	13.7	8	2	0.2
5	13.8	4	6	0.6
6	14.0	3	7	0.7
7	14.1	1	9	0.9
	14.4	0	10	1

As already mentioned above, the target criterion is the best fit of the quantile of the normal distribution to the data which is attained at a step position x_{exp} for the Heaviside function and a standard deviation s of

$$x_{exp} = 13.8034, \quad s = 0.2791$$

The goodness of fit may be assessed from figure 1.

An application of the "equal cumulative probabilities" criterion (using a frequency polygon for interpolating between the sampling points) results in an x_{exp} of 13.7834. For the standard deviation (33% quantile) one will obtain left- and right-hand side values which differ slightly one from the other

$$s_{left} = 0.2833, \quad s_{right} = 0.2067$$

and as the geometric mean an s of 0.2480.

The thus obtained values for x_{exp} are satisfactorily close to each other and do not exhibit any significant difference (the latter is as low as 7% of the corresponding uncertainty). There is also no statistical significant difference between the calculated standard deviations, and s_{left} coincides excellently with the s obtained from the normal distribution approach.

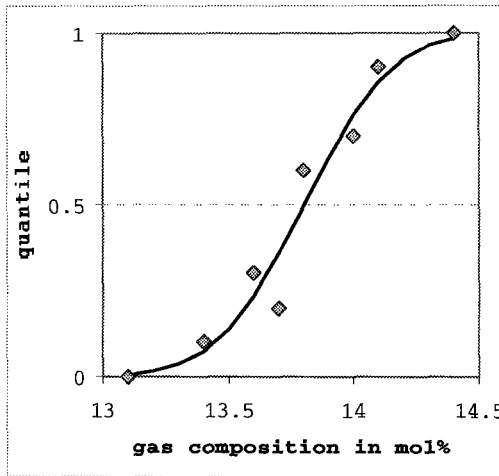


Fig. 1: Fit of the quantile of the normal distribution to experimental data.

Undoubtedly, one might argue whether asymmetric limits for the statistical scatter might or might not be an intrinsic feature of such semi-qualitative testing procedures. Since there are no widely accepted rules for the handling of asymmetric statistical boundaries, for the time being the normal distribution approach will be the method of choice which automatically generates symmetric standard deviations and confidence intervals. Time will show whether the introduction of the Appendix to the GUM dealing with Monte Carlo simulation (a method which quite often generates asymmetric boundaries) will change the above described situation.

Now, the position and the statistical boundaries of the step of the Heaviside function are determined. Figure 2 displays the function together with the scatter interval at the 67% confidence level. Figure 2 also displays the frequency polygon used for interpolation purposes between the sampling points, and the left- and right-hand side cumulated probabilities.

The confidence interval which can easily be calculated from a multiplication by the t factor for the corresponding number of degrees of freedom and the one-sided problem as being $CI(95\%) = 0.5300$ (alternatively: expansion with $k =$

2) covers in a sensible way the major part, but not all of the fuzzyness region of the test procedure.

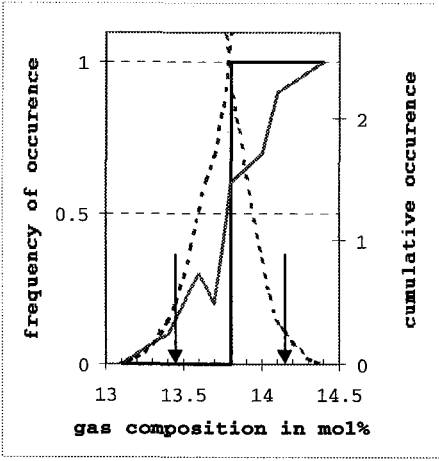


Fig. 2: Heaviside step function (black bold line) at the position x_{exp} , interpolated frequency of occurrence (grey bold line), cumulative occurrence (black dotted line, right scale), and statistical boundaries (black arrows) at the 67% confidence level.

In a last step the uncertainties of the parameter x (mixture composition) must be taken into account. The standard prEN 1839 requires that these uncertainties shall not exceed an absolute value of 0.2 mol% for all gas compositions in the composition range above 2 mol%. This absolute value is assigned to all sampling points of the experiment, and for uncertainty calculation each value x_i in the table is now subsequently reduced and increased by a value of one half of the uncertainty. The minimisation problem for the quantile is now solved with respect to only the x_{exp} (i.e. with fixed s) for each of the thus obtained, altered data sets.

For the third sampling point the corresponding line in the data table will be altered first to

#	x mol%	failure no ignition	success ignition	p
3	13.5	7	3	0.3

resulting in an $x_{exp} = 13.7948$, and subsequently to

#	x mol%	failure no ignition	success ignition	p
3	13.7	7	3	0.3

yielding an x_{exp} of 13.8217. The contribution of the third sampling point to the total uncertainty can now be calculated according to

$$u(x_{exp}: x_3) = 13.8217 - 13.7948 = 0.0269.$$

After geometric summation of the contributions from all sampling points and combination with the fuzzyness-of-assessment uncertainty, one will obtain a combined total uncertainty of $u_c(x_{exp}) = 0.3433$. As expected, the main contributors are sampling points 5 and 6.

4. Conclusions

The approach presented here is a possible way to tackle the problem of uncertainty assessment in semi-qualitative testing. However, several details and questions as raised above need further investigation and clarification.

Acknowledgments

The authors want to acknowledge the valuable contributions and helpful discussions of Dr K. Holtappels and Dr M. Malow, both with the BAM department "Chemical safety engineering", who raised the problem and provided some of the experimental data.

References

1. *ISO/IEC 17025:1999* General requirements for the competence of testing and calibration laboratories. ISO Geneva.
2. *ILAC-G17:2002* Introducing the Concept of Uncertainty of Measurement in Testing in Association with the Application of the Standard ISO/IEC 17025.
3. *ISO Guide to the Expression of Uncertainty in Measurement*, ISO Geneva, 1995.
4. *QUAM:2000.P1*, Quantifying Uncertainty in Analytical Measurement, EURACHEM/CITAC Guide.
5. *ISO 6143:2001*. ISO Geneva.

PROCESSING THE COHERENT ANOMALIES OF DIGITALIZED SURFACES IN WAVELET DOMAIN *

P. CIARLINI

*Istituto per le Applicazioni del Calcolo,
Viale del Policlinico 137,
00161 Rome, Italy
E-mail: ciarlini@iac.rm.cnr.it*

M. L. LO CASCIO

*Dipartimento di Matematica, Università di Messina,
Salita Sperone 31,
98166 Messina, Italy
E-mail: locascio@dipmat.unime.it*

Coherent anomalies can affect the quality of digitized images. In this paper the method, which was developed for their removal in images from old movies, is proposed in the contest of a metrological application in nanotechnology, where accurate thickness measurements of special coatings are obtained after an image processing. The method constructs a piece-wise spline approximation to restore the corrupted data in the low-pass filtered version of the digitized image obtained by a wavelet decomposition. An invisible reconstruction of the damaged area are obtained, since the type of the morphology in suitable domains is preserved. Simple statistical tests are able to automatically recognize the morphology type and to construct the appropriate approximation. The calibration study that have been performed for identify the test thresholds is described in the paper. It uses simulated images with different types of noise and of scratches. The benefit of the method adopted to locally preprocess the metrological image before acquiring the measurements is finally discussed.

1. Introduction

The digitalized images of nanotechnology surfaces can be affected by coherent anomalies, both coming from the surface itself and arising from the digitilizing process. The metrological application in nanotechnology

*Work partially funded under EU Project SofTools_MetroNet Contract n.G6RT-CT-2001-05061

is described in [7]: the preparation of the cross-section of the specimen causes distortion of the border between different materials; thickness measurements are obtained by measuring the distance between borders in the digitized image. A possible disalignment during the scanning process can cause the white vertical defects visible in figure 1.



Figure 1. Vertical anomalies in a metrological image from nanotechnology: in the left side the second white line crosses the lower border (in the digitized image it occurs at column 33)

Another example of vertical anomalies occurs in digitized images from old movie. These anomalies are caused by abrasion due to the contact of the film material with the mechanical parts of the projector. The vertical scratches are physically characterized by the partial loss of the film emulsion in a strip, with sidelobes of material; they are present in a image sequence and are seen by the audience as bright or dark lines. The anomalies appear as lines extending throughout almost the entire image, usually in the same position and with constant width.

In this paper, we propose to apply the strategy developed to remove the scratches in old movie images [2], to the digitized image of the metrological application. The method is able to preserve the image information content in the wavelet domain while removing the artifacts produced by the scanning operation. Therefore it should be applied before the accurate measuring of the coating thickness in the digitized image. In the following the linear anomaly will be named *scratch* as in the example of the old movie image.

A digitized image I is associated to a matrix $F = ((f_{ij})) \in \mathbb{R}^{(N_1, N_3)}$. The elements f_{ij} are the sampled values of the intensity function $f(x, y)$ at

the pixel (i, j) and they represent the Grey-level values of the image in a Grey-level scale related to the accuracy of the acquisition instrument. The function f is supposed in $L^2(\mathbb{R}^2)$, that is, a signal with finite energy. A scratch causes incorrect values in a certain number (say 5-10) of consecutive columns of the matrix F . The distortion, $l(\cdot, \cdot)$, is assumed to be additive. Therefore the sampled intensity function can be written as

$$f(x, y) = f^*(x, y) + l(x, y) + e(x, y). \quad (1)$$

where f^* is the intensity function of the cleaned image and e is a random gaussian noise.

A scratch removal can be formulated as a corrupted data problem in $L^2(\mathbb{R}^2)$ assuming in the damaged columns some information to be still available and the features similar to the ones present in the adjacent columns of F to be reconstructed in order to recover the information content $f^* + e$. The 1D approximation or interpolation methods, applied to each row of F , or 2D approximation methods in a strip of F , often provide a “visible” restoration, since they only partially deal with the high frequencies: the brightness discontinuity can be a feature of the original signal that can overlap the high frequency noise in the damaged area. In [1], [2] a new approach based on wavelet decomposition of F was used to separate the low frequency content from the high frequency details in order to appropriately treat the vertical features. The authors have used different wavelet bases and different approximation strategies applied to the matrices from the wavelet decomposition of F .

In the next section the three-step method developed in [2] is briefly described. To preserve the spatial correlation it adopts the strategy of choosing the approximation spaces according to the features that are present in a small area of the image adjacent to the scratch. The recognition of the features is performed by simple statistical tests applied in small areas. In Section 3 the calibration of the thresholds, used in the tests, and their robustness in presence of noisy images are discussed. In Section 4 the pre-processing of the metrological image by the three-step method is described and its benefits are shown. In the Appendix a short mathematical description of the bidimensional multiresolution analysis and of the filtering process are reported.

2. The three-step method for scratch removal

We assume the scratch on the image I be in the vertical direction, starting from the column k_s and with constant width w . The method operates

in three steps, namely a wavelet decomposition, an approximation phase and the wavelet reconstruction. Some mathematical backgrounds on the wavelet decomposition and reconstruction of the first and third step are given in the Appendix A. Here only a brief summary of the approximation step is reported.

The first step produces the low-pass filtered version of the original image I and the images containing the vertical, horizontal and diagonal details of I : they are indicated with the matrices A , V , H , R , respectively. Let us assume $A \in \mathbb{R}^{(256,256)}$.

According to the high characterization of the scratch in the vertical direction, the information relating to the linear defect is confined in A and V . Thus the matrices H and R of the horizontal and diagonal details are not processed, while in the matrices A and V only the columns corresponding to the scratch area will be corrected by using also the not corrupted information in the adjacent columns. The correction of the columns in V is obtained by applying a median filter.

In A , the data fitting procedure is complex. The parameters of the scratch, namely the initial column k_1 and the width w' , are computed from k_s , w and the length of the low pass filter, used in the first step. For example, $w = 5$ and the filters "bior3.1" (MATLAB notation) imply $w' = 3$.

The basic idea is now to replace the corrupted data in the columns of A with the values of a suitable bidimensional approximating function, say s , defined in the domain

$$D = [k_1 - w', k_1 + 2w' - 1] \times [1, 256].$$

It has horizontal dimension equal to $3w'$ and is centered on the scratch. The function s must be able to preserve the spatial correlation and the features, present in a neighborhood of the scratch. This requirement suggested to use a piece-wise function whose pieces are constructed according to the local morphology of the image. They are defined [2] in the following sub-domains of constant vertical dimension $d = 5$:

$$D_j = [k_1 - w', k_1 + 2w' - 1] \times [5(j - 1), 5j], \quad j = 1, 2, \dots$$

The sub-domains are classified into two *types*, *homogeneous*, and *not homogeneous*: the first type is characterized by small variations of the Grey-level values, otherwise in not homogeneous domains some edges are present. The criteria to recognize the absence/presence of features (edges) are discussed in the next section.

The 2D data fitting procedure of the second step constructs each piece of s as a tensor product of splines of the same order, namely a bilinear spline in homogeneous sub-domains, a bicubic spline, otherwise. In each one of the two cases, the optimal knot displacement is identified by means of an heuristic procedure described in [2] to solve the non linear minimization problem.

3. Calibration of the tests for feature recognition

Let us emphasize that the exact recognition of the type of a sub-domain D_j enables the method to use the appropriate spline space to get an invisible restoration, since the special feature in D_j is preserved.

The recognition method is based on simple first order statistics associated to a minor M of a matrix, namely the *standard deviation* $\sigma(M)$, the *mean value* $\mu(M)$ and the *median* $m(M)$.

Each D_j is divided into three equal parts with a horizontal dimension equal to w' , Σ_j^L , Σ_j , Σ_j^R , where Σ_j corresponds to the central corrupted columns, say Σ_j^L and Σ_j^R are the left and right neighborhoods, respectively. The following minors of the matrix $A = ((a_{ij}))$ are defined

$$\mathcal{L}_j = ((a_{il})), i, l \in \Sigma_j^L \quad (2)$$

$$\mathcal{C}_j = ((a_{il})), i, l \in \Sigma_j \quad (3)$$

$$\mathcal{R}_j = ((a_{il})), i, l \in \Sigma_j^R. \quad (4)$$

We have observed that the presence of an edge in the central block Σ_j corresponds to a significant increase of $\sigma(\mathcal{C}_j)$ with respect to $\sigma(\mathcal{C}_{j-1})$, while in homogeneous domains the mean and the median value do not differ significantly in the left and right parts, Σ_j^L and Σ_j^R , containing uncorrupted data. The following quantities are defined:

$$\delta_j^{(l)} = |\mu(\mathcal{L}_j) - m(\mathcal{L}_j)|, \quad \delta_j^{(r)} = |\mu(\mathcal{R}_j) - m(\mathcal{R}_j)|.$$

The plots in figure 2 show the δ differences related to the three minors in Eqs.(2- 4) for increasing value of j (they pertain to the image in figure 3). The behavior of the plots show that suitable thresholds can be identified to classify the two types of domains. We have chosen the threshold values $\omega_1, \omega_2, \omega_3$ independently of the image, in order to apply the method to several different images.

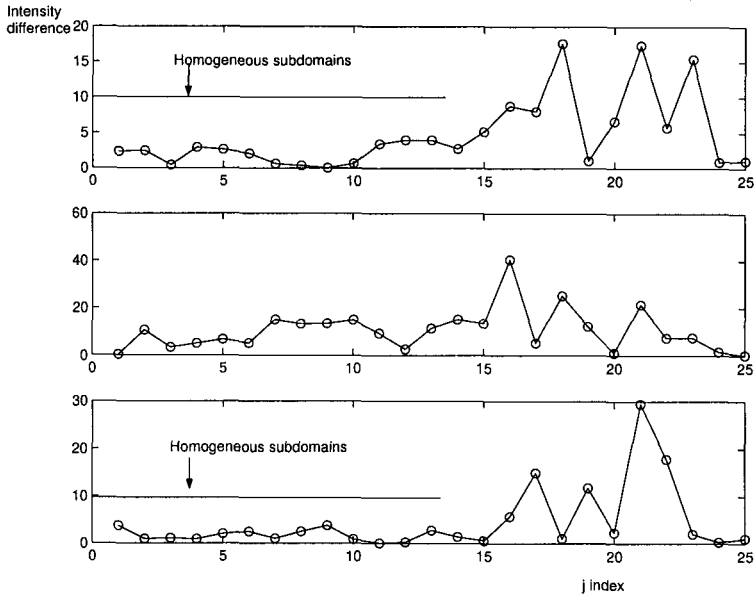


Figure 2. Plots of the difference between the mean and the median values at increasing j -index in Σ_j^L (top), Σ_j (middle), Σ_j^R (bottom)



Figure 3. KNIGHT image in Kokaram CD: domain segmentation of the scratch vertical region. In the column on the right side the results of the recognition method are shown (1 = homogeneous, 2 = not-homogeneous domain)

Hence we have defined a domain D_j to be *non homogeneous* if at least one of the following tests is verified:

$$\sigma(\mathcal{C}_j) - \sigma(\mathcal{C}_{j-1}) > \omega_1 \quad (5)$$

$$\delta_j^{(l)} > \omega_2 \text{ AND } \delta_j^{(r)} > \omega_2 \quad (6)$$

$$|\delta_j^{(l)} - \delta_j^{(r)}| > \omega_3. \quad (7)$$

The thresholds have been calibrated on a large class of test images, with 256 Grey-levels representation, from old movie (Kokaram [6] CD) and also from MATLAB images data bank.

To every image used in the calibration study, we have added simulated scratches in order to be able to compare the original unscratched image and the restored one. The additive scratch model in Eq. (1) is used with the function l deduced by the cross-section function in [9]. For fixed k_s , w , first column and width of the scratch, in any pixel (i, j) , $i = k_s, \dots, k_s + w - 1$, $j = 1, \dots, N_2$ of the original I the systematic effect is given by:

$$l(i, j) = 0.5^{|i-x_c|} \lambda \cos \frac{3\pi(i - x_c - rd(j))}{2w}, \quad x_c = k_s + \frac{w-1}{2} \quad (8)$$

where λ controls the intensity of the simulated scratch, and $rd(\cdot)$ is a random generator.

To assess the ability of the feature recognition tests, two criteria were used:

- the SNR (Signal-to-Noise Ratio) between the unscratched image and the restored one (with values f_{il}^*):

$$SNR(F) = \frac{\sum_{il} f_{il}^2}{\sum_{il} [f_{il} - f_{il}^*]^2} \quad (9)$$

- the good quality at a visual inspection (in the following named "eye norm").

In image analysis it is often necessary to add the "quality" criterion to the quantitative one, since the maximization of the SNR only does not always assure an invisible restoration. Indeed, the SNR is an overall criterion that may not reveal a defect of systematic type, as the human eye can.

The calibration study provided the following thresholds: $\omega_1 = 36$, $\omega_2 = 5$, $\omega_3 = 7$. When the Grey levels of the image belong to a different scale, the corresponding thresholds can be obtained by a scaling operation.

In figure 3 the vertical lines define the domain D and its subdivision in Σ_j^L , Σ_j and Σ_j^R , the horizontal lines separate the subdomains D_j . The results of the tests, reported in the column at right in the figure, show that the type of the D_j containing edges has been well classified.

Let's now discuss the robustness of the identified values of the thresholds of the tests in presence of highly noisy data via a simulation study. We have analysed the test images having simulated scratches using Eq. (8) for different values of the parameters. For three images in MATLAB database and with $w = 5$, the SNR values are reported in the second column of

Table 1. We have generated several noisy images using MATLAB routine "imnoise": white additive Gaussian noise with variance in $[0, 1e-3; 0, 1e-1]$ and Speckle multiplicative noise with variance in $[0, 2e-1; 0, 4e-1]$ are considered. Table 1 reports the SNR values for three test images and four types of noise. We underline that in every case the eye norm revealed a good restoration in spite of the bad values of the SNR function. An example of a detail in the Lena image is given in figure 4 where Gaussian noise with $\sigma^2 = 0,005$ was added. It must be noted that for values of the variance



Figure 4. Lena image with a simulated scratch in the eye zone and Gaussian noise with $\sigma^2 = 0,005$ (left), the restored image by the three-step method(right)

greater than the ones in the Table 1, an higher blurring effect was visible, therefore in the damaged region the domains with systematic features are always recognized as homogeneous, thus obtaining a visible restoration due to the presence of noise.

4. The metrological case-study in nanotechnology

In [7] an accurate measuring procedure for the thickness of special thin coatings of different magnetic materials was discussed. The preparation of the cross-section of the material and the misalignment of the scanning process by a magnetic sensor have produced noisy data in the border between the different material layers (for example the vertical scratch most visible is in the left side of figure 1). The task of the metrologist was to determine the precise contour exceeding fixed magnetic values using the

Table 1. Images with a simulated scratch ($w = 5$) and added noise

SNR (db)		$N(0; 0, 001)$	$N(0; 0, 005)$	$N(0; 0, 1e - 3)$	$S(0,02)$	$S(0,04)$
LENA	23,59	23,22	20,91	16,68	18,09	17,09
TRUCK	23,98	19,80	16,12	15,44	17,37	15,53
BABOON	19,16	19,00	17,27	16,00	16,55	14,62

digitized image (Grey-level visualization). The applied mathematical procedure was a three level wavelet decomposition with Haar basis to separate the horizontal features, and, at each level, a compression was also performed by neglecting the vertical and diagonal details: they do not contribute to the knowledge of the border position that has a horizontal direction. The metrologist measured the distance between two horizontal boundaries (at the same Grey-level), by means of a visual inspection of the processed digitized image. This mathematical procedure emphasizes the horizontal contours to get the thickness measurements in the Grey-level representation. However, neglecting every vertical detail might cause misleading information in predicting a reasonable position of the border when it crosses the vertical anomaly.

We now consider the anomaly that occurs in the left side of the image at $k_s = 33$ that has constant width $w = 5$. To get the measurements using the digitized image (4096 Grey-level scale), we need to accurately compute the horizontal contours: we apply the three-step method, before the multi-level wavelet analysis with compression, developed in the PTB Laboratory. The benefit due to our approximation step is to preserve the special feature of the lower border that also extends in the vertical direction near column 33. Indeed, some vertical details of the border position, can be recovered, which otherwise would be completely lost in the compression operation.

Table 2 shows the comparison of the results obtained in processing the image only by the method [7] and the ones obtained by adding the preprocessing with our method. The acquired measures of the borders, Y_u and Y_l are comparable. The main result is that with our preprocessing the thickness ΔY at level 0 (see last row in Table 2) can be measured since the upper border has been recovered in a more accurate way. Moreover, a similar benefit was obtained when comparing the results after only one level of the Haar decomposition and compression. Therefore we can suggest the use of our preprocessing and only one level of the Haar decomposition with compression, thus saving computing time.

Table 2. Thickness (ΔY) in the scratch area ($k_s = 33$) at 5 magnetic levels: columns 2-4 report the values obtained without preprocessing; columns 5-7 the values obtained with our preprocessing.

Level	Y_u upper border	Y_l lower border	ΔY	Y_u	Y_l	ΔY
4	168,0	104,1	63,9	165,0	104,2	60,8
3	172,5	98,4	74,2	172,4	98,4	74,1
2	176,5	94,9	81,7	176,5	94,9	81,7
1	185,0	92,7	92,3	184,0	92,7	91,3
0	-	92,4	-	193,0	92,4	100,6

5. Conclusions

This paper has addressed the problem of linear defect removal in images. The special treatment of these systematic anomalies affecting the image quality is useful when the digitized image must be used to obtain accurate measurements, such as in the discussed application in nanotechnology. Here the three-step method, developed for scratch removal in old movie images, was applied as an image preprocessing before the measuring procedure.

The novelty of the method is a biorthogonal wavelet decomposition coupled with the construction of appropriate spline approximations to restore the damaged low-pass filtered area. The good performance in reconstructing the local features that can be present in the image is due to the ability of three statistical tests to recognize the specific features in subdomains near the damaged area.

The simulation study, performed to calibrate the thresholds in the recognition tests, has showed the removal method in wavelet domains to be robust also in presence of noisy images. Therefore in the metrological image, the local features characterizing the borders crossing a vertical defect were recovered providing accurate measures. Moreover, an accurate determination of the thickness for a bigger number of Grey-levels have been obtained saving computing time.

Appendix A. Bidimensional multiresolution analysis and the filtering process

The first/third steps of the method in [2] for scratch removal consists of a wavelet decomposition/reconstruction in $L^2(\mathbb{R})$, where a multiresolution analysis [5] (MRA) has been defined. A multiresolution in $L^2(\mathbb{R}^2)$ can be easily obtained as tensor product of one dimensional multiresolution

analysis. More precisely let

$$\varphi(x) = \sum_{k \in \mathbb{Z}} h_k \varphi(2x - k), \quad h_k \in \mathbb{R}. \quad (\text{A.1})$$

be a *refinable function* in $L^2(\mathbb{R})$, that is the solution of the refinement equation. The mask $\mathbf{h} = \{h_k\}$ is assumed real and *finitely supported*, that is only a finite number of h_k are different from zero: consequently the *symbol* associated to φ

$$h(z) = \sum_{k \in \mathbb{Z}} h_k z^k$$

is a Laurent polynomial and the function φ is compactly supported.

A basic assumption on φ to get a MRA is that the integer translates, $\varphi(\cdot - k)$, form a Riesz basis: this hypothesis allows to define a *wavelet function*

$$\psi(x) = \sum_{k \in L} g_k \varphi(2x - k), \quad g_k \in \mathbb{R} \quad (\text{A.2})$$

with a finitely supported mask $\mathbf{g} = \{g_k\}$ and the subspaces of $L^2(\mathbb{R})$ giving the MRA

$$\begin{aligned} V_j &= \overline{\text{span}\{2^{j/2}\varphi(2^j x - k) =: \varphi_{jk}(x)\}} \\ W_j &= \overline{\text{span}\{2^{j/2}\psi(2^j x - k) =: \psi_{jk}(x)\}}. \end{aligned}$$

Then a *bidimensional multiresolution analysis* (2DMRA) in $L^2(\mathbb{R}^2)$ is obtained considering the spaces [5]

$$\begin{aligned} \mathbf{V}_j &= V_j \otimes V_j, \quad \mathbf{W}_j^h = V_j \otimes W_j \\ \mathbf{W}_j^v &= W_j \otimes V_j, \quad \mathbf{W}_j^d = W_j \otimes W_j \end{aligned}$$

A function $f \in \mathbf{V}_{j+1}$ can be decomposed by means the projections into the scaling space \mathbf{V}_j and the remainder spaces \mathbf{W}_j

$$f(x, y) = f_j(x, y) + r_j^h(x, y) + r_j^v(x, y) + r_j^d(x, y) := f_j(x, y) + R_j(x, y). \quad (\text{A.3})$$

By the inclusion property of the \mathbf{V}_j the decomposition can be iterated at consecutive levels $j-1, \dots, j-L$

$$f(x, y) = f_{j-L}(x, y) + R_{j-L}(x, y) + R_{j-L+1}(x, y) + \dots + R_j(x, y) \quad (\text{A.4})$$

where any $R_k = r_k^h + r_k^v + r_k^d$ is given by three components.

Vice versa, given the decomposition Eq. (A.4) it is possible to reconstruct the original function f with an inverse procedure.

When the bases are not orthogonal a new pair $(\tilde{\varphi}, \tilde{\psi})$ of a refinable and a wavelet function, which generates a MRA $\{\tilde{V}_j, \tilde{W}_j\}$, is considered, in order to compute the generalized Fourier expansion of a given function [5]. The pairs (φ, ψ) and $(\tilde{\varphi}, \tilde{\psi})$ are *dual* or *biorthogonal*, in the sense that

$$\begin{aligned} \langle \varphi_{j,n}, \tilde{\varphi}_{j',n'} \rangle &= \delta_{j,j'} \delta_{n,n'}, \quad \langle \varphi_{j,n}, \tilde{\psi}_{j',n'} \rangle = 0 \\ \langle \psi_{j,n}, \tilde{\psi}_{j',n'} \rangle &= \delta_{j,j'} \delta_{n,n'}, \quad \langle \psi_{j,n}, \tilde{\varphi}_{j',n'} \rangle = 0. \end{aligned}$$

Let be $\tilde{\mathbf{h}} = \{\tilde{h}_k\}$, $\tilde{\mathbf{g}} = \{\tilde{g}_k\}$ the real and finitely supported masks of $\tilde{\varphi}$ and $\tilde{\psi}$. The above conditions can be expressed in terms of the symbols associated to the refinable and wavelet functions [3], that is

$$h(z)\tilde{h}(z^{-1}) + h(-z)\tilde{h}(-z^{-1}) = 1, \quad z = e^{-i\omega} \quad (\text{A.5})$$

(i imaginary unit) with h, \tilde{h} normalized as $h(1) = \tilde{h}(1) = 1$ and $g(z) = z^\alpha \tilde{h}(-z)$, $\tilde{g}(z) = z^\alpha h(-z)$, with α odd. Both the wavelet decomposition and the reconstruction can be interpreted in terms of signal filtering. In fact it is known that the sequences $\{h_k\}$ and $\{\tilde{h}_k\}$ can be assumed as digital low-pass FIR filters and $\{g_k\}$ and $\{\tilde{g}_k\}$ as digital high-pass FIR filters. In this context $h(z), \tilde{h}(z), g(z), \tilde{g}(z)$ are the transfer function of the filters and Eq.(A.5) is the perfect reconstruction condition. Moreover the projection operations on the spaces of the MRA reproduce the classical sub-band coding [8].

Then the filtered versions of a given image obtained by a sub-band coding are respectively the projections into the refinable space and into the W type spaces (see Eq. (A.3)). They can be viewed as function in $L^2(\mathbb{R}^2)$.

The digitized values of the four images are obtained by convolution decimation operators having a tensor product structure

$$\begin{aligned} P_a f(u, v) &= \sum_{j,k \in \mathbb{Z}} f(j, k) h_{2n-j+1} h_{2n'-k+1} \\ P_h f(u, v) &= \sum_{j,k \in \mathbb{Z}} f(j, k) h_{2n-j+1} g_{2n'-k+1} \\ P_v f(u, v) &= \sum_{j,k \in \mathbb{Z}} f(j, k) g_{2n-j+1} h_{2n'-k+1} \\ P_d f(u, v) &= \sum_{j,k \in \mathbb{Z}} f(j, k) g_{2n-j+1} h_{2n'-k+1}. \end{aligned}$$

The dual filters $\tilde{\mathbf{h}} = \{\tilde{h}_k\}$, $\tilde{\mathbf{g}} = \{\tilde{g}_k\}$ allow to introduce the analogous adjoint operator $P_a^* P_h^* P_v^* P_d^*$ involved in the reconstruction operation.

Indeed, it is the tensor product structure that allows to separate the high frequency details according to the horizontal, vertical and diagonal directions.

In conclusion, a digital image represented by a matrix F can be decomposed into four matrices by the operators P_λ : the matrix from the operator P_a , say A , is the low-pass filtered version of F ; the matrices from the operators P_v and P_h , V and H , contain the vertical and horizontal details of the image; the matrix from the operator P_d , R , contains the diagonal details.

In the three-step method [2] the biorthogonal B-spline were suggested since they show a better visual behavior in the restoration than the Haar basis, generally adopted. To keep the computational cost low, bases with minimum support are to be preferred, for example the "bior2.2" or "bior3.1" filters (MATLAB notation).

References

1. T. Bretschneider, O. Kao and P.J. Bones, *Proc. of Image and Vision Computing New Zealand*, 37 (2000).
2. P. Ciarlini, G. Costanzo, M.L. Lo Cascio, *Proc. Wavelets & Splines, S. Petersburg 2003* (submitted).
3. A. Cohen, R.D. Ryan, *Wavelets and multiscale signal processing*, Chapman-Hall, London (1995).
4. C. de Boor, *A practical guide to spline*. Springer-Verlag, Berlin (1998).
5. I. Daubechies, *Ten lectures*. SIAM, Philadelphia (1992).
6. A. Kokaram, *Motion Picture Restoration*. Springer Berlin (1998).
7. H. Moeck and K. Herrmann, In *Advanced Mathematical and Tools in Metrology V*, Series on Advances in Mathematics for applied sciences, 57 Ciarlini, Cox, Filipe, Pavese, Richter Eds, Singapore, World Scientific, 270 (2001).
8. G. Strang, T. Nguyen, *Wavelets and filter banks*, Wellesley-Cambridge Press (1997).
9. D. Vitulano, V. Bruni and P. Ciarlini, Line scratches detection on digital images: an energy based model. *10-th WSCG International Conference 2002 Praga*, 10, 2, 477 (2002).

LEAST SQUARES ADJUSTMENT IN THE PRESENCE OF DISCREPANT DATA*

M G COX, A B FORBES, J L FLOWERS AND P M HARRIS

National Physical Laboratory, Teddington, TW11 0LW, UK

Email: maurice.cox@npl.co.uk, alistair.forbes@npl.co.uk

jeff.flowers@npl.co.uk, peter.harris@npl.co.uk

Least squares methods provide a flexible and efficient approach for analyzing metrology data. Given a model, measurement data values and their associated uncertainty matrix, it is possible to define a least squares analysis method that gives the measurement data values the appropriate ‘degree of belief’ as specified by the uncertainty matrix. Least squares methods also provide, through χ^2 values and related concepts, a measure of the conformity of the model, data and input uncertainty matrix with each other. If there is conformity, we have confidence in the parameter estimates and their associated uncertainty matrix. If there is nonconformity, we seek methods of modifying the input information so that conformity can be achieved. For example, a linear response may be replaced by a quadratic response, data that has been incorrectly recorded can be replaced by improved values, or the input uncertainty matrix can be adjusted. In this paper, we look at a number of approaches to achieving conformity in which the main element to be adjusted is the input uncertainty matrix. These approaches include the well-known Birge procedure. In particular, we consider the natural extensions of least squares methods to maximum likelihood methods and show how these more general approaches can provide a flexible route to achieving conformity. This work was undertaken as part of the Software Support for Metrology and Quantum Metrology programmes, funded by the United Kingdom Department of Trade and Industry.

1. Introduction

Least squares analysis methods are widely used in metrology and are justified for a large number of applications on the basis of both statistical theory and practical experience. Given a model specified by parameters \mathbf{a} , measurement data values and their associated uncertainty matrix, it is possible to define a least squares analysis (LSA) method that gives the measurement data values the appropriate ‘degree of belief’ as specified by the uncertainty

*Work partially funded under EU SofTools_MetroNet Contract N. G6RT-CT-2001-05061

matrix. LSA also provides a concise method of evaluating the uncertainty matrix associated with estimates of the fitted parameters.

Through χ^2 values and related concepts, it is possible to give a measure of the conformity (or consistency) of the model, data and input uncertainty matrix with each other. If there is conformity, we have confidence in the parameter estimates and their associated uncertainty matrix. If there is nonconformity, we seek methods of modifying the input information so that conformity can be achieved. Nonconformity indicates that the ensemble of input information consisting of both model assumptions and observations represents an unlikely state and that parameter estimates and their associated uncertainties based on this input information may therefore be invalid. By reassessing the input information in the light of the nonconformity, we may be able to develop an alternative, more self-consistent interpretation of the input information or to quantify the changes in the information that would be required to bring about conformity.

In section 2, we summarize the main elements of LSA along with solution algorithms. In section 3 we look at a number of approaches to achieving conformity in which the main element to be adjusted is the input uncertainty matrix. These approaches include the well-known Birge procedure. In particular, we consider the natural extensions of least squares methods to maximum likelihood methods and show how these more general approaches can provide a flexible route to achieving conformity. We also indicate in this section how these procedures can be implemented in a numerically stable way. We illustrate these procedures on a number of applications in section 4. Our concluding remarks are given in section 5.

2. Least squares analysis

Suppose we have a linear model in which the responses $\boldsymbol{\eta} = (\eta_1, \dots, \eta_m)^T$ of a system are determined by parameters $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)^T$ through a fixed, known $m \times n$ matrix C so that $\boldsymbol{\eta} = C\boldsymbol{\alpha}$. We assume that $m \geq n$ and C is of rank n . We wish to determine estimates of the parameters $\boldsymbol{\alpha}$ from measurements of responses η_i . Suppose the measurement model is

$$\mathbf{Y} = \boldsymbol{\eta} + \mathbf{E} \quad (1)$$

with $E(\mathbf{E}) = \mathbf{0}$, $\text{Var}(\mathbf{E}) = V$, and that we observe measurements \mathbf{y} of \mathbf{Y} . If the uncertainty matrix V has full rank, the least squares estimate \mathbf{a} of $\boldsymbol{\alpha}$, given \mathbf{y} , minimizes $\chi^2 = (\mathbf{y} - C\mathbf{a})^T V^{-1}(\mathbf{y} - C\mathbf{a})$. For the case $V = I$, the identity matrix, $\mathbf{a} = C^\dagger \mathbf{y}$, where $C^\dagger = (C^T C)^{-1} C^T$ is the *pseudo-inverse* of C . While the estimate \mathbf{a} can be calculated thus in terms of C , a better

approach^{3,7} is to factorize C as $C = QR$, where Q is an $m \times m$ orthogonal matrix with $Q^T Q = QQ^T = I$ and R is an $m \times n$ upper-triangular matrix. Writing

$$C = QR = [Q_1 \ Q_2] \begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix} = Q_1 R_1,$$

we see that $C^\dagger = R_1^{-1} Q_1^T$ and \mathbf{a} hence solves the $n \times n$ triangular system $R_1 \mathbf{a} = Q_1^T \mathbf{y}$. Here Q_1 (Q_2) is the submatrix of Q comprising the first n (last $m - n$) columns.

For a general (full rank) uncertainty matrix V with a factorization $V = LL^T$ (cf. section 3.4), where L is an $m \times m$ matrix, also necessarily full rank, the least squares estimate is given by

$$\mathbf{a} = \tilde{C}^\dagger \tilde{\mathbf{y}}, \quad \tilde{C} = L^{-1} C, \quad \tilde{\mathbf{y}} = L^{-1} \mathbf{y},$$

where \tilde{C}^\dagger is the pseudo-inverse of \tilde{C} . For well conditioned V and L , this approach is satisfactory. However, if L is poorly conditioned the formation and use of \tilde{C} , etc., can be expected to introduce numerical errors. The *generalized QR factorization*^{4,10,11} approach avoids this potential numerical instability. Suppose $V = LL^T$, where L is $m \times p$. (Often $p = m$ but the approach applies in the more general case.) The estimate \mathbf{a} can be found by solving

$$\min_{\mathbf{a}, \mathbf{z}} \mathbf{z}^T \mathbf{z} \quad \text{subject to constraints} \quad \mathbf{y} = C\mathbf{a} + L\mathbf{z}. \quad (2)$$

Note that if L is invertible,

$$\mathbf{z} = L^{-1}(\mathbf{y} - C\mathbf{a}), \quad \mathbf{z}^T \mathbf{z} = (\mathbf{y} - C\mathbf{a})^T V^{-1}(\mathbf{y} - C\mathbf{a}).$$

We factorize $C = QR$ and $Q^T L = TU$ where R and T are upper-triangular and Q and U are orthogonal. Multiplying the constraints by Q^T , we have

$$\begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \tilde{\mathbf{y}}_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix} \mathbf{a} + \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{z}}_1 \\ \tilde{\mathbf{z}}_2 \end{bmatrix}, \quad (3)$$

where $\tilde{\mathbf{y}} = Q^T \mathbf{y}$, and $\tilde{\mathbf{z}} = U\mathbf{z}$. From the second set of equations, $\tilde{\mathbf{z}}_2$ must satisfy $\tilde{\mathbf{y}}_2 = T_{22} \tilde{\mathbf{z}}_2$. Given any $\tilde{\mathbf{z}}_1$, the first set of equations is satisfied if $R_1 \mathbf{a} = \tilde{\mathbf{y}}_1 - T_{11} \tilde{\mathbf{z}}_1 - T_{12} \tilde{\mathbf{z}}_2$. We choose $\tilde{\mathbf{z}}_1 = \mathbf{0}$ in order to minimize $\mathbf{z}^T \mathbf{z} = \tilde{\mathbf{z}}^T \tilde{\mathbf{z}} = \tilde{\mathbf{z}}_1^T \tilde{\mathbf{z}}_1 + \tilde{\mathbf{z}}_2^T \tilde{\mathbf{z}}_2$, so that \mathbf{a} solves $R_1 \mathbf{a} = \tilde{\mathbf{y}}_1 - T_{12} \tilde{\mathbf{z}}_2$. Public-domain library software for solving (2) and, more generally, computing generalized QR factorizations is available¹¹.

2.1. Uncertainty matrix associated with the least squares estimate

For fixed C and \mathbf{Y} an m -vector of random variables, the equation $\mathbf{A} = C^\dagger \mathbf{Y}$ defines the n -vector \mathbf{A} of random variables as linear combinations of \mathbf{Y} . Taking expectations, we have

$$E(\mathbf{A}) = E(C^\dagger \mathbf{Y}) = C^\dagger \boldsymbol{\eta} = C^\dagger C \boldsymbol{\alpha} = \boldsymbol{\alpha},$$

and

$$V_{\mathbf{A}} = \text{Var}(\mathbf{A}) = C^\dagger V(C^\dagger)^T = (C^T V^{-1} C)^{-1}.$$

2.2. Conformity

Consider first the case where $\text{Var}(\mathbf{Y}) = I$. If $X_i \sim N(0, 1)$, $i = 1, \dots, m$, then $\sum_{i=1}^m X_i^2$ has a χ_m^2 distribution with $E(\chi_m^2) = m$ and $\text{Var}(\chi_m^2) = 2m$. Let \mathbf{R} be the random vector of residuals so that

$$\mathbf{R} = \mathbf{Y} - C\mathbf{A} = \mathbf{Y} - CC^\dagger \mathbf{Y} = (I - CC^\dagger)\mathbf{Y}.$$

If $C = Q_1 R_1$, then $CC^\dagger = Q_1 Q_1^T$ and $I - Q_1 Q_1^T = Q_2 Q_2^T$, so that

$$S^2 = \mathbf{R}^T \mathbf{R} = (Q_2^T \mathbf{Y})^T Q_2^T \mathbf{Y}.$$

Now Q is orthogonal so setting $\tilde{\mathbf{Y}} = Q\mathbf{Y}$ we have $\text{Var}(\tilde{\mathbf{Y}}) = I$ also. Therefore, $S^2 = \sum_{i=n+1}^m \tilde{Y}_i^2$ is a sum of squares of $m - n$ normal variates and has a χ^2 distribution with $\nu = m - n$ degrees of freedom, with $E(S^2) = \nu$, $\text{Var}(S^2) = 2\nu$. In general, given a least squares estimate $\mathbf{a} = \mathbf{a}(\mathbf{y}, V)$ associated with a conforming model and data, we expect

$$\chi^2 = (\mathbf{y} - C\mathbf{a})^T V^{-1} (\mathbf{y} - C\mathbf{a}) \approx \nu.$$

If χ^2 is far from ν (relative to $(2\nu)^{1/2}$) we may wish to reassess \mathbf{y} , V , \mathbf{a} and $V_{\mathbf{A}}$.

2.3. Example applications

Interlaboratory comparison exercises. In an interlaboratory comparison exercise, one or more artefacts are measured by a number of laboratories, with the i th laboratory recording a measured value $y_{i,k}$ for the k th measurand α_k and an associated standard uncertainty $u_{i,k}$. From $\{y_{i,k}\}$ and $\{u_{i,k}\}$ we generally wish to determine estimates y_k of the measurands α_k and their associated uncertainties u_k . The analysis of such data is often complicated

by the fact that the dispersion of the $\{y_{i,k}\}$ cannot be accounted for easily by the stated uncertainties $\{u_{i,k}\}$.

Adjustment of the fundamental constants. There are many fundamental constants of physics, a small number of which such as the speed of light in a vacuum have exact assigned values, the others being estimated from measurements provided a large number of experiments. Generally, each experiment provides an estimate y_i of $\eta_i = c_i(\alpha)$, a quantity related to some subset of the fundamental constants α , and an associated standard uncertainty u_i . From this ensemble of measurements⁵, estimates of the constants are found by solving a large nonlinear least squares problem. The last such adjustment exercise was undertaken in 1999⁸ and took account of all the relevant data available up to 31 December 1998 relating to over 50 constants. In such exercises, the conformity of the model and input data is of great concern since, due to the complex functional inter-relationships, an invalid input could have a significant effect on a number of parameter estimates.

3. Adjustment procedures

In this section, we consider ways of adjusting the input information \mathbf{y} and V in order to achieve conformity of the least squares solution. We regard the least squares estimate $\mathbf{a} = \mathbf{a}(\mathbf{y}, V)$ as a function of the measurements \mathbf{y} and the input uncertainty matrix V . Similarly,

$$\chi^2 = \chi^2(\mathbf{y}, V) = (\mathbf{y} - C\mathbf{a})^T V^{-1}(\mathbf{y} - C\mathbf{a})$$

is also a function of \mathbf{y} and V . We look for adjustments $(\mathbf{y}, V) \longrightarrow (\hat{\mathbf{y}}, \hat{V})$ so that $\hat{\chi}^2 = \chi^2(\hat{\mathbf{y}}, \hat{V}) = \nu$ (or acceptably close to ν in terms of the distribution χ_ν^2). Rather than allowing complete freedom to adjust \mathbf{y} and V , we may consider a modest number of degrees of freedom parametrized by $\lambda = (\lambda_1, \dots, \lambda_p)^T$ so that $\hat{\mathbf{y}} = \hat{\mathbf{y}}(\lambda)$, $\hat{V} = \hat{V}(\lambda)$ and $\chi^2 = \chi^2(\lambda)$. We seek adjustments that, in some sense, make minimal changes to \mathbf{y} and V in order to bring about conformity. Our definition of minimal will reflect our belief in the input estimates \mathbf{y} and V and the underlying model.

From a computational point of view we would like some measure $\|\bullet\|$ so that the adjustment problem can be posed as

$$\min_{\mathbf{a}, \lambda} M(\lambda) = \|(\hat{\mathbf{y}}(\lambda), \hat{V}(\lambda)) - (\mathbf{y}, V)\| \quad (4)$$

subject to the constraints

$$\left. \begin{aligned} C^T \widehat{V}^{-1}(\lambda) C \mathbf{a} &= C^T \widehat{V}^{-1}(\lambda) \hat{\mathbf{y}}(\lambda), \\ (\hat{\mathbf{y}}(\lambda) - C \mathbf{a})^T \widehat{V}^{-1}(\lambda) (\hat{\mathbf{y}}(\lambda) - C \mathbf{a}) &= \nu. \end{aligned} \right\} \quad (5)$$

The first constraint states that \mathbf{a} is the least squares solution associated with $(\hat{\mathbf{y}}(\lambda), \widehat{V}(\lambda))$, while the second is the conformity constraint. Below, we usually assume that only V is to be adjusted, but all the algorithms can be adapted for the more general case.

3.1. One-parameter adjustment

In the case where we allow only one degree for freedom, parametrized by λ , for adjustment, the conformity constraint specifies λ completely (if we ignore the possibility of multiple solutions) and no measure $\|\bullet\|$ is required.

*Birge procedure*¹. In the Birge procedure,

$$\widehat{V}(\lambda) = \lambda V, \quad (6)$$

representing a re-scaling of the input uncertainty matrix. If \mathbf{r} is the vector of residuals associated with \mathbf{y} and V , then $\lambda = \nu / (\mathbf{r}^T V^{-1} \mathbf{r})$. This procedure reflects a belief that the input uncertainty matrix is known only up to a scale factor and is often used in the case where $V = \sigma^2 I$, where σ is unknown. The estimate \mathbf{a} remains unchanged.

*Hidden independent random effect*¹³. In this procedure, $\widehat{V}(\lambda) = V + \lambda I$, corresponding to the belief that each observation was subject to hidden, independent random effects with zero mean and unknown variance λ . The associated measurement model is

$$\mathbf{Y} = \boldsymbol{\eta} + \mathbf{E} + \mathbf{F}, \quad E(\mathbf{F}) = \mathbf{0}, \quad \text{Var}(\mathbf{F}) = \lambda I. \quad (7)$$

Parameters \mathbf{a} and λ are specified by the nonlinear equations (5). We note that if V has eigen-decomposition⁷ $V = UDU^T$, with U orthogonal and D diagonal, then $V + \lambda I = U(D + \lambda I)U^T$ and

$$[V + \lambda I]^{-1} = U(D + \lambda I)^{-1}U^T.$$

In general $\mathbf{a}(\lambda) \neq \mathbf{a}(0) \equiv \mathbf{a}$, but, for the case $V = \sigma^2 I$, this procedure is equivalent to the Birge procedure.

*Hidden correlated random effects*¹². Let \mathbf{r} be the vector of residuals associated with \mathbf{y} and V and set $\widehat{V}(\lambda) = V + \lambda \mathbf{r} \mathbf{r}^T$. Doing so corresponds to the measurement model

$$\mathbf{Y} = \boldsymbol{\eta} + \mathbf{E} + F \mathbf{r}, \quad E(F) = 0, \quad \text{Var}(F) = \lambda. \quad (8)$$

In this case, $\mathbf{a}(\lambda) = \mathbf{a}$.

All of the above procedures have the form $\widehat{V}(\lambda) = V_0 + \lambda V_1$. We note that, from the generalized eigenvalue decomposition⁷ for the pair (V_0, V_1) , we can find a nonsingular matrix X such that $D_0 = X^T V_0 X$ and $D_1 = X^T V_1 X$ are diagonal matrices. It follows that

$$\widehat{V}^{-1}(\lambda) = X(D_0 + \lambda D_1)^{-1} X^T.$$

3.2. Maximum likelihood approaches to adjustment

For adjustment procedures involving parameters $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_p)^T$, $p > 1$, it is necessary to define a measure $M(\boldsymbol{\lambda})$ as in (4). Rather than define *ad hoc* measures, we consider ones having a probabilistic justification. We make the further assumption about the model (1) that $\mathbf{E} \sim N(\mathbf{0}, V(\boldsymbol{\lambda}))$. Then, given data $\mathbf{y} \in \mathbf{Y}$, the probability $p(\mathbf{y}|\mathbf{a}, \boldsymbol{\lambda})$ that the data arose from parameters $(\mathbf{a}, \boldsymbol{\lambda})$ is given by

$$p(\mathbf{y}|\mathbf{a}, \boldsymbol{\lambda}) = \frac{1}{|2\pi\widehat{V}(\boldsymbol{\lambda})|^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{y} - C\mathbf{a})^T \widehat{V}^{-1}(\boldsymbol{\lambda})(\mathbf{y} - C\mathbf{a}) \right\},$$

and the log likelihood $L(\mathbf{a}, \boldsymbol{\lambda}|\mathbf{y})$ by

$$-L(\mathbf{a}, \boldsymbol{\lambda}|\mathbf{y}) = \frac{1}{2} \log |2\pi\widehat{V}(\boldsymbol{\lambda})| + \frac{1}{2}(\mathbf{y} - C\mathbf{a})^T \widehat{V}^{-1}(\boldsymbol{\lambda})(\mathbf{y} - C\mathbf{a}).$$

The maximum likelihood (ML) estimates of \mathbf{a} and $\boldsymbol{\lambda}$ are determined by maximizing the probability $p(\mathbf{y}|\mathbf{a}, \boldsymbol{\lambda})$ which is equivalent to minimizing $-L(\mathbf{a}, \boldsymbol{\lambda}|\mathbf{y})$ with respect to \mathbf{a} and $\boldsymbol{\lambda}$.

From its definition, an ML approach finds the most likely explanation of the data. To the extent that the quantity $\chi^2 - \nu$ is also a measure of the likelihood of the data, we generally expect ML methods to produce a χ^2 value that is as reasonable as the adjustment procedure will allow. For example, if $\widehat{V} = \lambda V$, the ML estimate is $\lambda_{ML} = \mathbf{r}^T \mathbf{r} / m$, whereas the conforming $\lambda_C = \mathbf{r}^T \mathbf{r} / (m - n)$. The more flexibility there is in the adjustment procedure, the ‘easier’ it is for an ML estimate to satisfy the conformity constraint. However, increasing the flexibility also means that the ML optimization problem is more likely to be ill-posed. For example, we can apply an ML approach for the case $\widehat{V}(\boldsymbol{\lambda}) = \lambda_1 V + \lambda_2 I$ where we allow both the possibility of re-scaling and hidden random effects. If V is close to a multiple of the identity matrix I , the corresponding estimation problem is poorly posed since the two parameters λ_1 and λ_2 specify essentially the same behaviour. In these situations, we can augment the log likelihood function with a regularization term $R(\boldsymbol{\lambda})$ in order to help resolve any such

ambiguity. Again, we would like the regularization term to encode, in a probabilistic way, our prior belief about the likely range of values for λ .

With this generalization, we can then calculate the most likely conforming solution by solving

$$\min_{\lambda, \mathbf{a}} \log |\hat{V}(\lambda)| + R(\lambda) \quad (9)$$

subject to the constraints (5).

3.3. Algorithms for diagonal uncertainty matrices

We now consider the case where the uncertainty matrix V is diagonal and the adjusted $\hat{V}(\lambda)$ is also constrained to be diagonal. Let $D = D(\lambda) = \hat{V}^{-1}(\lambda)$. If $\tilde{\mathbf{y}} = D^{1/2}\mathbf{y}$, $\tilde{C} = D^{1/2}C$, and $\tilde{C} = \tilde{Q}\tilde{R}$ is its QR factorization, then \mathbf{a} solves the linear least squares problem $\tilde{C}\mathbf{a} = \tilde{\mathbf{y}}$. That \mathbf{a} is constrained to be the least squares solution associated $\hat{V}(\lambda)$ can be stated as

$$C^T D(\lambda) C \mathbf{a} = C^T D(\lambda) \mathbf{y} \quad (10)$$

and defines $\mathbf{a} = \mathbf{a}(\lambda)$ implicitly as a function of λ . If \dot{D} and $\dot{\mathbf{a}}$ denote the derivative of D and \mathbf{a} with respect to one of the parameters λ_q , then

$$C^T D C \dot{\mathbf{a}} = C^T \dot{D} (\mathbf{y} - C \mathbf{a}),$$

defines $\dot{\mathbf{a}}$ in terms of \dot{D} . We note that this system of linear equations can be solved by solving two triangular systems involving \tilde{R} .

The associated log likelihood L for a data set \mathbf{y} is given by

$$-L(\mathbf{a}, \lambda | \mathbf{y}) = \frac{1}{2} \sum_i \log 2\pi \hat{v}_{ii}(\lambda) + \frac{1}{2} \sum_i \frac{1}{\hat{v}_{ii}(\lambda)} (y_i - \mathbf{c}_i^T \mathbf{a})^2.$$

Consider the case where $\hat{v}_{ii}(\lambda) = \lambda_i$, $i = 1, \dots, m$, and our belief in the input V is encoded as

$$\log v_{ii} \sim N(\log \lambda_i, (\log \rho)^2), \quad \rho \geq 1.$$

Roughly, this says that we are 95% certain that $\lambda_i/\rho^2 \leq v_{ii} \leq \lambda_i\rho^2$. (We might prefer to use, for example, a gamma distribution instead of a log normal distribution to represent our confidence in V , but the general approach would be essentially the same.) An appropriate regularization term is therefore

$$R(\lambda) = \frac{1}{(\log \rho)^2} \sum_{i=1}^m (\log v_{ii} - \log \lambda_i)^2.$$

The generalized ML estimate is found by solving

$$\min_{\mathbf{a}, \boldsymbol{\lambda}} \frac{1}{2} \left\{ \sum_i \log \lambda_i + \sum_i \frac{1}{\lambda_i} (y_i - \mathbf{c}_i^T \mathbf{a})^2 + R(\boldsymbol{\lambda}) \right\}.$$

(If we believed that v_{ii} and y_i were observations associated with independent random variables, then this objective function is essentially $-L(\mathbf{a}, \boldsymbol{\lambda} | \mathbf{y}, V)$, where L is the log likelihood function. In practice, v_{ii} and y_i are quite likely to be jointly distributed in which case the joint distribution would be used in the definition of the log likelihood function.) The most likely conforming estimate is found by solving

$$\min_{\boldsymbol{\lambda}} \frac{1}{2} \left\{ \sum_i \log \lambda_i + R(\boldsymbol{\lambda}) \right\}, \quad (11)$$

subject to the conformity constraint

$$\sum_i \frac{1}{\lambda_i} (y_i - \mathbf{c}_i^T \mathbf{a}(\boldsymbol{\lambda}))^2 = \nu,$$

regarding $\mathbf{a} = \mathbf{a}(\boldsymbol{\lambda})$ as defined implicitly by (10). If required we can impose additional equality or inequality constraints on λ_i . For example, if we believe that no v_{ii} is an underestimate we can impose the constraints

$$\lambda_i \geq v_{ii}, \quad i = 1, \dots, m. \quad (12)$$

3.4. Algorithms for general uncertainty matrices

For the case of general uncertainty matrices, we wish to use a generalized QR factorization approach to avoid potential numerical instabilities associated with forming and using the inverses of matrices. We also need to be able to calculate the gradient of the objective and constraint functions in (9) in order to employ efficient constrained optimization algorithms^{6,9}.

Any symmetric positive definite matrix V has a Cholesky factorization of the form $V = LL^T$, where L is a lower triangular matrix. The matrix V can similarly be factored as $V = TT^T$ where T is an upper-triangular matrix. These factorizations can be computed in a numerically stable way using a simple algorithm. For example, consider the factorization

$$TT^T = \begin{bmatrix} T_{11} & \mathbf{t}_{12} \\ & t_{22} \end{bmatrix} \begin{bmatrix} T_{11}^T & \\ \mathbf{t}_{12}^T & t_{22} \end{bmatrix} = \begin{bmatrix} V_{11} & \mathbf{v}_{12} \\ \mathbf{v}_{12}^T & v_{22} \end{bmatrix} = V.$$

Equating terms we have $t_{22}^2 = v_{22}$, $t_{22}\mathbf{t}_{12} = \mathbf{v}_{12}$ and $T_{11}T_{11}^T = V_{11} - \mathbf{t}_{12}\mathbf{t}_{12}^T$. The problem is now reduced to finding the factorization of the modified

submatrix $V_{11} - \mathbf{t}_{12}\mathbf{t}_{12}^T$. The complete factorization can be achieved by repeating this step:

- I Set $T(i, j) = V(i, j)$, for all $i \geq j$.
- II For $k = n : -1 : 1$, set $T(k, k) = T(k, k)^{1/2}$ and

$$T(1 : k - 1, k) = T(1 : k - 1, k)/T(k, k).$$

- II.j For $j = k - 1 : -1 : 1$, set

$$T(1 : j, j) = T(1 : j, j) - T(1 : j, k)T(j, k).$$

Suppose now that $V = V(\boldsymbol{\lambda}) = T(\boldsymbol{\lambda})T^T(\boldsymbol{\lambda})$, and $\dot{V} = \frac{\partial V}{\partial \lambda_q}$ and $\dot{T} = \frac{\partial T}{\partial \lambda_q}$. The matrix \dot{T} satisfies $\dot{T}T^T + T\dot{T}^T = \dot{V}$ and can be determined by differentiating the algorithm² to compute T :

- I Set $\dot{T}(i, j) = \dot{V}(i, j)$, for all $i \geq j$.
- II For $k = n : -1 : 1$, set $\dot{T}(k, k) = \dot{T}(k, k)/(2T(k, k))$ and

$$\dot{T}(1 : k - 1, k) = (\dot{T}(1 : k - 1, k) - \dot{T}(k, k)T(1 : k - 1, k))/T(k, k).$$

- II.j For $j = k - 1 : -1 : 1$, set

$$\dot{T}(1 : j, j) = \dot{T}(1 : j, j) - T(1 : j, k)\dot{T}(j, k) - \dot{T}(1 : j, k)T(j, k).$$

This algorithm can be easily vectorized to compute all the partial derivatives of $T(\boldsymbol{\lambda})$ simultaneously.

Turning to the general problem (4-5), let $C = QR$, set $\tilde{\mathbf{y}}(\boldsymbol{\lambda}) = Q^T \hat{\mathbf{y}}(\boldsymbol{\lambda})$ and factorize $Q\hat{V}(\boldsymbol{\lambda})Q^T = T(\boldsymbol{\lambda})T^T(\boldsymbol{\lambda})$. Corresponding to (3) we have

$$\begin{bmatrix} \tilde{\mathbf{y}}_1(\boldsymbol{\lambda}) \\ \tilde{\mathbf{y}}_2(\boldsymbol{\lambda}) \end{bmatrix} = \begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix} \mathbf{a} + \begin{bmatrix} T_{11}(\boldsymbol{\lambda}) & T_{12}(\boldsymbol{\lambda}) \\ & T_{22}(\boldsymbol{\lambda}) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{z}}_1 \\ \tilde{\mathbf{z}}_2 \end{bmatrix}, \quad (13)$$

and from the analysis above we know that if $\tilde{\mathbf{y}}_2(\boldsymbol{\lambda}) = T_{22}(\boldsymbol{\lambda})\tilde{\mathbf{z}}_2$ then the \mathbf{a} that solves $R_1 \mathbf{a} = \tilde{\mathbf{y}}_1(\boldsymbol{\lambda}) - T_{12}(\boldsymbol{\lambda})\tilde{\mathbf{z}}_2$ is the least squares solution associated with $\hat{\mathbf{y}}(\boldsymbol{\lambda})$ and $\hat{V}(\boldsymbol{\lambda})$. Moreover the solution \mathbf{a} will satisfy the χ^2 constraint if $\tilde{\mathbf{z}}_2^T \tilde{\mathbf{z}}_2 = \nu$. Since this latter constraint does not involve \mathbf{a} we can reformulate (4-5) as

$$\min_{\boldsymbol{\lambda}, \tilde{\mathbf{z}}_2} M(\boldsymbol{\lambda}) \quad \text{subject to} \quad \tilde{\mathbf{y}}_2(\boldsymbol{\lambda}) = T_{22}(\boldsymbol{\lambda})\tilde{\mathbf{z}}_2, \quad \tilde{\mathbf{z}}_2^T \tilde{\mathbf{z}}_2 = \nu.$$

In this formulation, there is no requirement to form the inverse of a matrix and the partial derivatives of the constraint equations are easily determined from the partial derivatives of $\mathbf{y}(\boldsymbol{\lambda})$, $V(\boldsymbol{\lambda})$ and its factorization.

4. Numerical examples

In this section we compare the behaviour of four adjustment procedures: A) Birge procedure (6), B) hidden independent effects (7), C) hidden correlated effects (8) and D) a maximum likelihood approach (11) incorporating constraints (12) to ensure that no input uncertainty is reduced. The latter procedure was implemented straightforwardly using the NAG⁹ constrained optimization routine E04UCF. Below $u_{ii} = v_{ii}^{1/2}$ and $\hat{u}_{ii} = \hat{v}_{ii}^{1/2}$, the input and adjusted standard uncertainties associated with the observation y_i .

4.1. Measurements of G

The first set of data concerns the ten measurements of the Newtonian constant of gravitation G considered by Weise and Wöger¹². Fig. 1 plots the estimates y_i and corresponding coverage intervals $y_i \pm 2u_{ii}$ and $y_i \pm 2\hat{u}_{ii}$ associated with the input (first interval) and three adjusted uncertainties corresponding to algorithms B, C and D. Observations 4 and 7 have relatively large input uncertainties. Observation 10 has a small input uncertainty but a value considerably different from the other measurements. In order to achieve conformity using procedure A, all the input uncertainties have to be multiplied by 22.6. Procedure B achieves conformance by large relative increases to all the uncertainties apart from those associated with observations 4 and 7. Procedure C has a more modest effect on all the uncertainties except that for observation 10, which is multiplied by over 20. Procedure C also introduces appreciable correlations (with many correlation coefficients over 0.9 in magnitude), which are not apparent in Fig. 1. The main effect of procedure D is to increase the uncertainty associated with observations 5 and 10 by factors of nearly four and over 30, respectively.

Another view of the behaviour of the adjustment procedures is given in Fig. 2 which plots the residuals $\mathbf{r} = \mathbf{y} - C\mathbf{a}$ and corresponding intervals $r_i \pm 2u_{ii}$ associated with the least squares fit to the input data and $\hat{r}_i \pm 2\hat{u}_{ii}$ associated with the fit for the three adjusted uncertainties corresponding to algorithms B, C and D. Only with procedure D are the residuals scattered around zero. For all the other procedures the fit appears skewed by the effect of observation 10. More recent measurements of G provide more evidence that observation 10 is discrepant.

4.2. Simple adjustment model

To investigate adjustment procedures for determining estimates of the physical constants we consider a much simpler model involving five (fictitious)

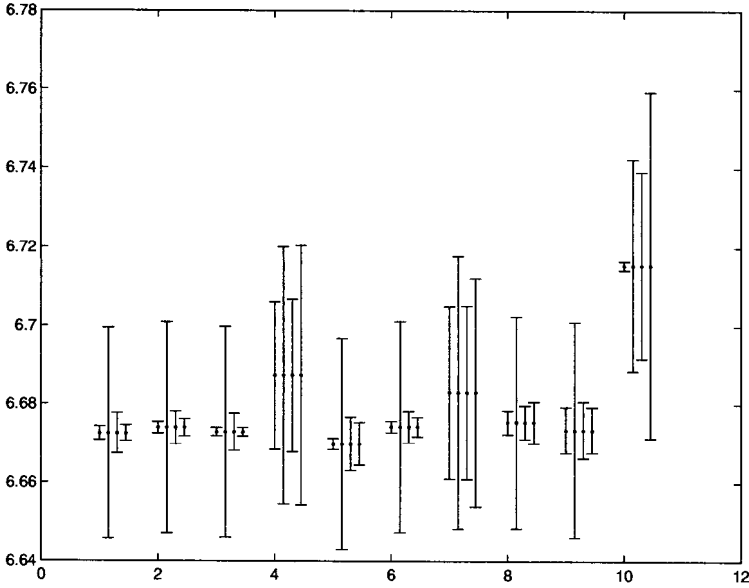


Figure 1. Input and adjusted uncertainties corresponding to procedures B, C, and D and for the 10 measurements of G .

physical constants $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ and $\beta = (\beta_1, \beta_2)^T$ and nine measurements with associated observation matrix

$$C = \begin{bmatrix} 100 & 100 & 5 & 0 & -5 \\ 100 & -100 & 0 & 5 & -5 \\ 100 & 0 & 100 & -5 & 5 \\ 100 & 0 & -100 & -5 & 5 \\ 0 & 100 & 100 & 5 & -5 \\ 0 & 100 & -100 & 5 & 5 \\ 5 & 0 & 0 & 100 & 100 \\ 0 & 0 & -5 & 100 & 100 \\ 0 & 5 & 0 & 100 & -100 \end{bmatrix}.$$

The first six (last three) observations contain strong information about α (β). Given $\gamma = (\alpha^T, \beta^T)^T$, we generate exact data according to $\eta = C\gamma$ and then simulate measurement data by adding perturbations $y = \eta + e$ according to some model. For example, we have generated exact η for $\gamma_j = 1$, $j = 1, \dots, 5$, added random Gaussian noise to η , $y_i = \eta_i + e_i$, $e_i \in N(0, \sigma_i^2)$, $\sigma_i \in [0.5, 3.0]$, and then subtracted 10 from y_2 , simulating a discrepant value. Fig. 3 shows the perturbations $y - \eta$ and corresponding intervals $y_i - \eta_i \pm 2u_{ii}$ and $y_i - \eta_i \pm 2\hat{u}_{ii}$ associated with the input (first

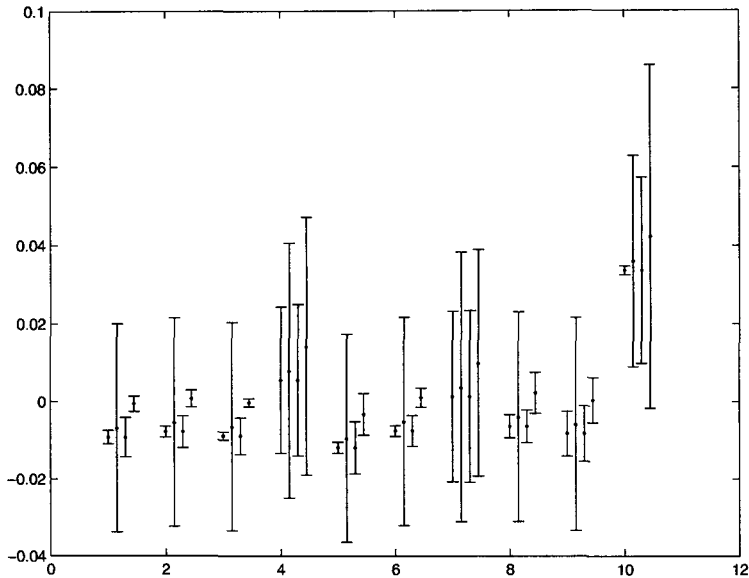


Figure 2. Residuals $\mathbf{r} = \mathbf{y} - \mathbf{C}\mathbf{a}$ along with error bars $r_i \pm 2u_{ii}$ associated with the least squares fit to the input data and $\hat{r}_i \pm 2\hat{u}_{ii}$ associated with the fit for the three adjusted uncertainties corresponding to algorithms B, C and D.

interval) and four adjusted uncertainties corresponding to algorithms A, B, C and D. Both procedures A and B result in a large increase in all the input uncertainties, with procedure B leading to a more uniform set of uncertainties. Procedure C has a different behaviour with significant increases in the uncertainties associated with observations 2, 3, and 5 and small or modest increases for the others. Procedure D produces a large increase in the uncertainty associated with the second (discrepant) observation and modest increases associated with observations 8 and 9.

The residuals $\mathbf{r} = \mathbf{y} - \mathbf{C}\mathbf{a}$ for the fits associated with procedures A, B and D are plotted in Fig. 4 along with the perturbations $y_i - \eta_i$. (Procedure C gives the same fit as procedure A.) By assigning a large uncertainty to observation 2, procedure D is able to tolerate a large residual. The residuals for D approximate well the perturbations $y_i - \eta_i$. The largest residual error associated with procedure A is for observation 5 and there is no indication that observation 2 is discrepant in any way. The behaviour of procedure B is somewhere between A and D.

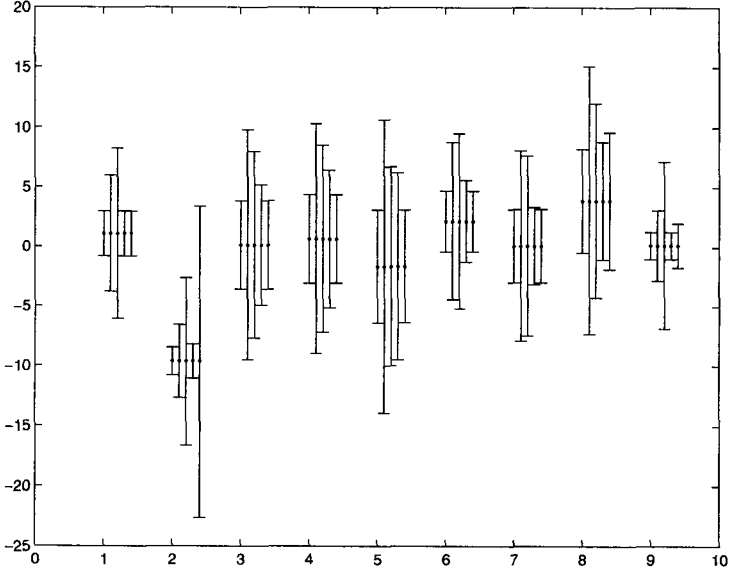


Figure 3. Input and adjusted uncertainties corresponding to procedures A, B, C and D for the nine observations with one discrepant data point.

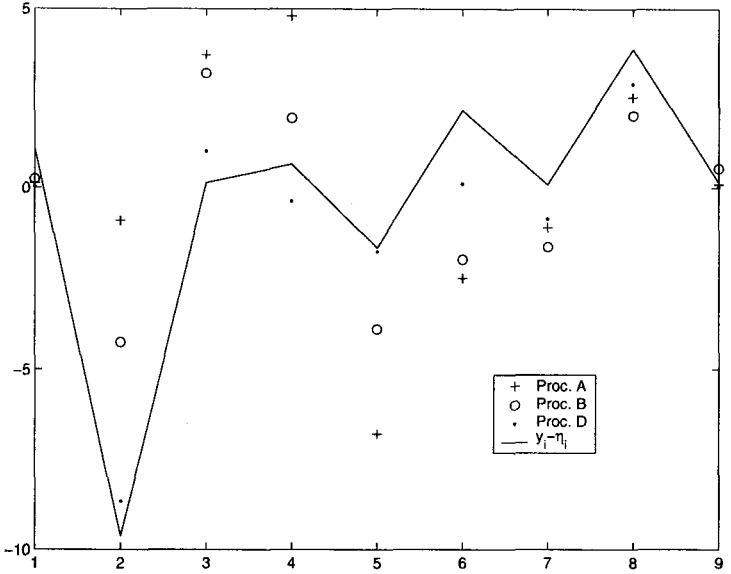


Figure 4. Residual errors associated with procedures A, B and D along with perturbations $y_i - \eta_i$ (joined by a line).

5. Concluding remarks

In this paper we have been concerned with procedures for adjustment of input information to achieve conformity of the associated least squares solution. We have provided a general framework for such procedures and described algorithms to bring about conformity. These algorithms avoid potential numerical instabilities by using appropriate matrix factorization approaches. We have been concerned to define procedures that have a probabilistic foundation rather than *ad hoc* methods. In particular, we have described adjustment methods based on maximum likelihood estimation that combine flexibility with an appropriate use of the available information.

Acknowledgments. The authors are grateful for the generous support from SofTools MetroNet and to attendees for many helpful discussions.

References

1. R. T. Birge. Probable values of the general physical constants. *Rev. Mod. Phys.*, 1:1–73, 1929.
2. R. Boudjemaa, M. G. Cox, A. B. Forbes, and P. M. Harris. Automatic differentiation and its applications to metrology. In *this volume*.
3. M. G. Cox, A. B. Forbes, and P. M. Harris. Software Support for Metrology Best Practice Guide 4: Modelling Discrete Data. Technical report, National Physical Laboratory, Teddington, 2000.
4. M. G. Cox, A. B. Forbes, P. M. Harris, and I. M. Smith. Classification and solution of regression problems for calibration. Technical Report CMSC 24/03, National Physical Laboratory, May 2003.
5. J. L. Flowers and B. W. Petley. Progress in our knowledge of the fundamental constants of physics. *Rep. Prog. in Phys.*, pages 1191–1246, 2001.
6. P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, London, 1981.
7. G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, third edition, 1996.
8. P. J. Mohr and B. N. Taylor. CODATA recommended values of the fundamental physical constants: 1998. *Rev. Mod. Phys.*, 72:351–495, 2000.
9. The Numerical Algorithms Group Limited, Oxford. *The NAG Fortran Library, Mark 20, Introductory Guide*, 2002. <http://www.nag.co.uk/>.
10. C. C. Paige. Fast numerically stable computations for generalized least squares problems. *SIAM J. Numer. Anal.*, 16:165–171, 1979.
11. SIAM, Philadelphia. *The LAPACK User's Guide*, third edition, 1999.
12. K. Weise and W. Woeger. Removing model and data non-conformity in measurement evaluation. *Measurement Science and Technology*, 11:1649–1658, 2000.
13. R. Willink. Statistical determination of a comparison reference value using hidden errors. *Metrologia*, 39:343–354, 2002.

HARMONIZATION OF CORRELATED CALIBRATION CURVES WITH AN APPLICATION TO THE ANALYSIS OF NATURAL GASES*

M. G. COX, S. KAMVISSIS, M. J. T. MILTON AND G. VARGHA

National Physical Laboratory

Queens Road, Teddington, Middlesex TW11 0LW, UK

E-mail: maurice.cox@npl.co.uk, sarantis.kamvissis@npl.co.uk,

martin.milton@npl.co.uk, gergely.vargha@npl.co.uk

During the analysis of natural gas by gas chromatography, the indicated response is affected systematically by varying environmental and instrumental conditions. Calibration data dispersion is attributed not only to random effects, but also to systematic measurement effects. The accuracy of the measurement results and consequently the usability of a calibration model are accordingly reduced. The model consists of a series of calibration curves, one for each component of the gas measured. When the systematic uncertainty component dominates, a correlated response behaviour between corresponding points on which the set of curves depends is observed. A model-based least-squares method is introduced that compensates for these effects. It incorporates correction parameters to account for the systematic uncertainty components, thus eliminating the correlation effects and reducing the calibration model uncertainty. Results are presented for calibration data modelled by straight-line functions. Generalizations are indicated.

1. Introduction

1.1. *Measurement process*

N samples, each from a different standard, that consist of the same q gas components, are injected in turn to a gas chromatograph (GC) (figure 1). A flow of carrier gas (e.g., helium) supports sample circulation through the column of the GC. During each run, corresponding to the analysis of one standard, the injected gas component molecules are naturally separated. The detector within the GC indicates any deviation from the carrier gas

*Work partially funded under EU project SofTools_MetroNet Contract N. G6RT-CT-2001-05061 and supported by the Valid Analytical Measurement and Software Support for Metrology programmes of the UK's Department of Trade and Industry.

condition. It provides spectral peaks, one for each gas component present (figure 2). The area under each peak relates to the molar fraction of gas component measured.

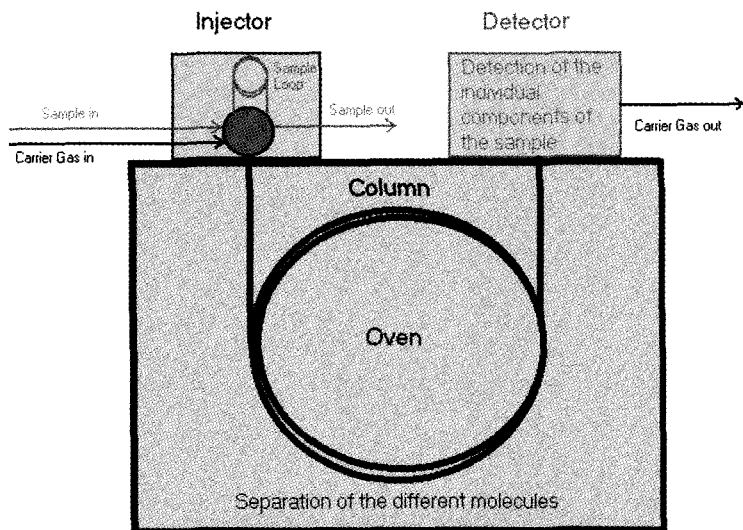


Figure 1. Schematic representation of the gas chromatograph.

The application of the Ideal Gas Law gives

$$n_{ij} = \frac{p_i V}{RT_i} x_{ij}, \quad (1)$$

where n_{ij} is the number of molecules of the j th gas component participating in the molecular interaction in terms of x_{ij} , the amount fraction (gravimetric concentration) of that gas component in the i th standard, V , the volume of the sample loop, R , the ideal gas constant, p_i , the ambient pressure, and T_i , the temperature during the i th measurement run. Expression (1) applies for $j = 1, \dots, q$, the indices of the gas components processed by the detector, where $q \leq Q$, the total number of gas components constituting the sample, and $i = 1, \dots, N$, the indices of the standards used.

x_{ij} remains fixed by definition, since it is the relative abundance of gas molecules while, if n_i is the total number of molecules injected to the GC

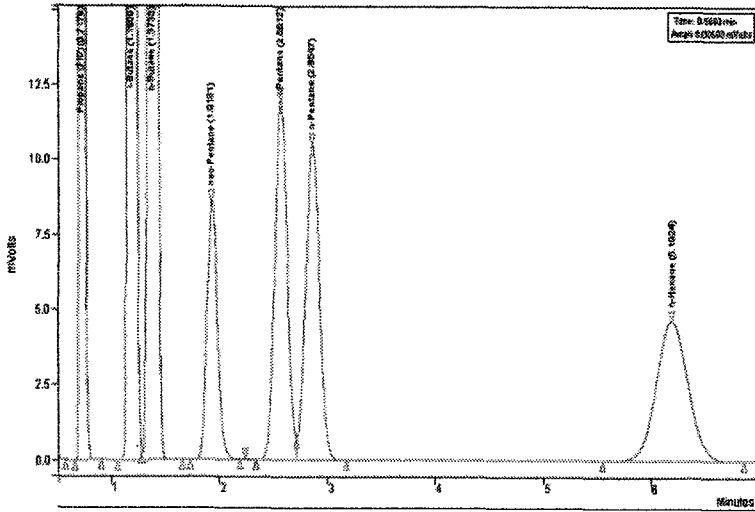


Figure 2. Spectral peaks provided by a gas chromatograph.

for the i th standard,

$$x_{ij} = \frac{n_{ij}}{n_i}, \quad n_i = \sum_{j=1}^Q n_{ij}.$$

Changes in the environmental and instrumental conditions (e.g., ambient pressure, temperature) between measurement runs affect the numbers n_{ij} of injected molecules, but not their concentration ratios:

$$\frac{n_{ij}}{n_{ik}} = \frac{n_{ij}/n_i}{n_{ik}/n_i} = \frac{x_{ij}}{x_{ik}}.$$

If f_j denotes the relative response factor of the detector, the theoretical area y_{ij} recorded by the detector for the j th gas component in the i th standard is, using expression (1),

$$y_{ij} = f_j n_{ij} = f_j \frac{p_i V}{T_i R} x_{ij}. \quad (2)$$

1.2. Experimental considerations

The changes in the environmental and instrumental conditions between runs that affect the n_{ij} accordingly influence in a systematic manner the indicated areas y_{ij} for all q measured gas components. Let ε_{ij} denote the random error, regarded as a signal from an unknown source, associated with

the measurement result y_{ij} . Let $u(v)$ denote the standard uncertainty associated with an estimate v of the value of a quantity. When the systematic uncertainty component dominates the data, i.e.,

$$u(p_i/T_i) \gg u(\varepsilon_{ij}), \quad (3)$$

a correlated response behaviour—termed *covariation*—is observed among the placement of the calibration curves for all q gas components participating in the i th experimental run. Conversely, when the random uncertainty component dominates the calibration data, the inequality (3) is reversed, and the indicated area values are randomly dispersed about the derived calibration curves.

The objective is to adjust the measurement results to eliminate the systematic effects. The adjusted calibration data would then deviate only randomly from the fitted model. Consideration is thus given to the application of multiplicative correction factors to the standards to improve the consistency of the indicated responses y_{ij} . The use of the corrected data set reduces the uncertainty associated with the calibration curves and achieves an improved simultaneous model fit. One purpose of this work is to quantify that reduction for actual data.

The calibration data consists of stimulus values x_{ij} and corresponding response measurements y_{ij} , $i = 1, \dots, N$, for each gas component (indexed by) $j = 1, \dots, q$. This data generally exhibits an underlying proportional behaviour (as in figure 3):

$$y_{ij} = f_j \frac{p_i}{T_i} \frac{V}{R} x_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, N. \quad (4)$$

Expression (4) constitutes the measurement model used.

Figure 4 shows a calibration data set free from covariation, where $u(y_{ij}) \approx u(\varepsilon_{ij})$.

Figure 5 shows an example of covariation between data points. Corresponding data points are influenced similarly, tending to be either 'high' or 'low' together. Following correction it is expected that $u(y_{ij}) \approx u(\varepsilon_{ij})$, i.e., that the uncertainty is essentially associated only with a random process.

The determination and use of straight-line calibration curves is considered in section 2. Numerical methods for constructing the calibration curve parameters are the subject of section 3. The uncertainties associated with the calibration curve parameters and with the molar fractions obtained from the use of the calibration curves are addressed in section 4. Results obtained using the approach are given in section 5. Possible extensions

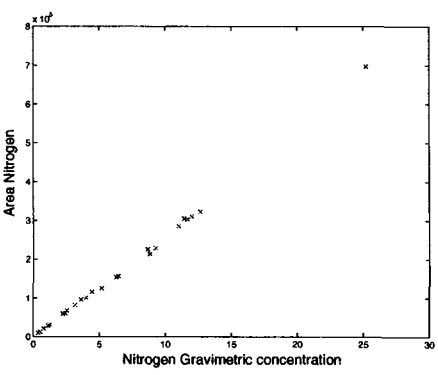


Figure 3. Typical calibration data, having an underlying straight-line behaviour.

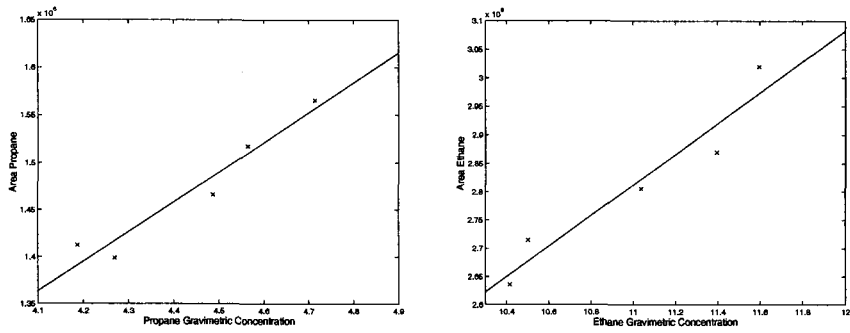


Figure 4. An example of data free from covariation.

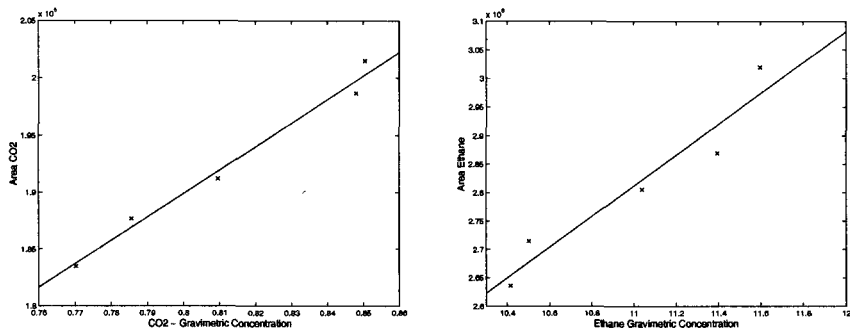


Figure 5. An example of covariation between data points

of the approach, to higher-order calibration curves and other uncertainty structures, are indicated in section 6.

2. Determination and use of the calibration curves

The determination of the calibration curves is considered first without accounting for systematic effects, and then the required corrections are introduced within a multiplicative model to derive improved calibration curves.

The calibration curves are then used (inversely) to obtain the molar fractions x_j^* , say, from values of response (area) y_j^* .

2.1. Straight-line model

In practice, the detector response may exhibit a small offset as well as the proportional effect implied by expression (4). The linear regression model equation ¹

$$y_{ij} = \widehat{\alpha}_j + \widehat{\beta}_j x_{ij} + \eta_{ij} \quad (5)$$

is therefore used to represent this behaviour, in which $\widehat{\alpha}_j$ and $\widehat{\beta}_j$ are the offset and gradient estimates when measuring the j th gas component, and the error processes η_{ij} comprise both random and systematic effects. The calibration curves are formed by solving ⁴

$$\min_{\alpha_j, \beta_j} \sum_{i=1}^N (y_{ij} - (\alpha_j + \beta_j x_{ij}))^2, \quad j = 1, \dots, q,$$

for each gas component separately, or, equivalently, simultaneously by solving

$$\min_{\alpha, \beta} \sum_{j=1}^q \sum_{i=1}^N (y_{ij} - (\alpha_j + \beta_j x_{ij}))^2$$

where $\alpha = (\alpha_1, \dots, \alpha_q)^T$ and $\beta = (\beta_1, \dots, \beta_q)^T$, giving parameter estimates $\widehat{\alpha} = (\widehat{\alpha}_1, \dots, \widehat{\alpha}_q)^T$ and $\widehat{\beta} = (\widehat{\beta}_1, \dots, \widehat{\beta}_q)^T$, and a resulting sum of squared deviations

$$S = \sum_{j=1}^q \sum_{i=1}^N (y_{ij} - (\widehat{\alpha}_j + \widehat{\beta}_j x_{ij}))^2. \quad (6)$$

2.2. Multiplicative model

In order to derive improved calibration curves, introduce multiplicative correction factors c_i so that

$$c_i \frac{p_i}{T_i} = \frac{p_0}{T_0}$$

for some ‘reference’ values p_0 and T_0 . Then, using expression (2),

$$y_{ij} = \frac{1}{c_i} y_{ij}^0, \quad y_{ij}^0 = f_j \frac{p_0}{T_0} \frac{V}{R} x_{ij} = c_i y_{ij},$$

where the y_{ij}^0 constitute corrected values of y_{ij} . The modified model assigns the degree c_i to which the fractions p_i/T_i deviate from the ideal constant fraction p_0/T_0 . Such corrections compensate for systematic measurement effects during the measurement process.

The linear regression model equation, including offsets as before, and regarding the random effects as following a normal distribution with mean value zero and variance σ_ε^2 , becomes

$$\widehat{c}_i y_{ij} = \widehat{\alpha}_j + \widehat{\beta}_j x_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim NID(0, \sigma_\varepsilon^2). \quad (7)$$

Calibration curves are then determined by solving

$$\min_{\alpha, \beta, c} \sum_{j=1}^q \sum_{i=1}^N (y_{ij} c_i - (\alpha_j + \beta_j x_{ij}))^2, \quad (8)$$

where $\mathbf{c} = (c_1, \dots, c_N)^T$. The resulting ‘corrected’ residual sum of squares is

$$S_{\text{corrected}} = \sum_{j=1}^q \sum_{i=1}^N (y_{ij} \widehat{c}_i - (\widehat{\alpha}_j + \widehat{\beta}_j x_{ij}))^2. \quad (9)$$

Separation into q independent residual sums of squares is not now possible as the presence of the corrections \mathbf{c} results in interaction effects among the terms in the outer summation. Estimation of model parameters is achieved by solving a system of $m = qN$ linear equations in $n = 2q + N$ unknowns α , β and \mathbf{c} to deliver $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\mathbf{c}} = (\widehat{c}_1, \dots, \widehat{c}_N)^T$.

There are only $2q + N - 1$ *mutually independent* unknowns since it is necessary to impose a relationship on \mathbf{c} (below). For a unique least-squares solution to be possible, there must be at least as many equations as independent unknowns, i.e., $qN \geq 2q + N - 1$. For the system to be over-determined, $qN \geq 2q + N$. In terms of designing the experiment, it is generally desirable to maximize the degree of over-determinacy, i.e.,

for $qN \gg 2q + N$, since doing so reduces the influence of the random effects on the solution.

The problem (8) is, however, ill-posed, since it has the trivial and physically meaningless solution $\hat{\alpha} = \hat{\beta} = \hat{c} = \mathbf{0}$, giving $S_{\text{corrected}} = 0$. To obtain a valid solution postulate that the corrections \mathbf{c} are realizations of a discrete random variable C with mean unity and probability mass function $^3 P(C = c_i) = 1/N$. Then, denoting expectation by E ,

$$1 = E(C) = \sum_{i=1}^N c_i P(C = c_i) = \sum_{i=1}^N c_i / N.$$

Thus, the following ‘resolving’ constraint is imposed on the solution:

$$\sum_{i=1}^N c_i = N. \quad (10)$$

2.3. Use of the calibration curves

Given area values y_j^* , the inverse use of the calibration curves yields estimates of the unknown molar fractions x_j^* :

$$x_j^* = \frac{y_j^* - \hat{a}_j}{\hat{\beta}_j} \quad (11)$$

3. Numerical methods

Express the measurements of the j th gas component as

$$\mathbf{y}_j = (y_{1j}, \dots, y_{Nj})^T, \quad j = 1, \dots, q,$$

and let

$$\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_q^T]^T$$

denote the complete set of qN measurements. Let the calibration parameters be expressed as

$$\mathbf{b} = [\mathbf{b}_1^T, \dots, \mathbf{b}_q^T]^T, \quad \mathbf{b}_j = (\alpha_j, \beta_j)^T.$$

For the linear regression model (5) with uncoupled pairs \mathbf{b}_j , $j = 1, \dots, q$, of calibration parameters,

$$X\mathbf{b} \simeq \mathbf{y} \quad \text{or} \quad X_j \mathbf{b}_j \simeq \mathbf{y}_j, \quad j = 1, \dots, q,$$

where

$$X = \text{diag}\{X_j\}_1^q, \quad X_j = \begin{bmatrix} 1 & x_{1j} \\ \vdots & \vdots \\ 1 & x_{Nj} \end{bmatrix}.$$

The least-squares solution is

$$\hat{\mathbf{b}} = (X^T X)^{-1} X^T \mathbf{z}, \quad \text{or} \quad \hat{\mathbf{b}}_j = (X_j^T X_j)^{-1} X_j^T \mathbf{z}_j, \quad j = 1, \dots, q.$$

For a *general* linear regression model $X\mathbf{b} \simeq \mathbf{y}$, where X is an $m \times n$ design matrix, $m \geq n$ and X has rank n , the solution can be obtained in a numerically stable way by using an orthogonal factorization $X = QR$ of X , where Q is orthogonal and R upper-triangular.¹⁰ R , the Cholesky factor of X , satisfies $R^T R = X^T X$. $\hat{\mathbf{b}}$ is then obtained by solving the triangular system $R\hat{\mathbf{b}} = Q^T \mathbf{y}$.^a This approach is used for all least-squares problems considered here.

For the linear regression model (7) with coupled pairs of calibration parameters,

$$[X - S] \mathbf{b}' \simeq \mathbf{0}, \quad (12)$$

where

$$S = (\text{diag}(\mathbf{y}_1), \dots, \text{diag}(\mathbf{y}_q))^T, \quad \mathbf{b}' = (\mathbf{b}^T, \mathbf{c}^T)^T.$$

The resolving constraint (10) can be incorporated through a Lagrange multiplier or, directly, by expressing c_N in terms of c_1, \dots, c_{N-1} . Using the latter approach, the system (12) becomes

$$X\mathbf{b}' \simeq \mathbf{z}, \quad X = \begin{bmatrix} X_1 & T_1 \\ \vdots & \vdots \\ X_q & T_q \end{bmatrix}, \quad T_j = \begin{bmatrix} -S'_j \\ y_{Nj} \mathbf{e}_{N-1}^T \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_q \end{bmatrix},$$

with $S'_j = \text{diag}(y_{1j}, \dots, y_{N-1,j})$, $\mathbf{z}_j = [0, \dots, 0, Ny_{Nj}]^T$ and \mathbf{e}_r a column vector of r ones. This system has a unique least-squares solution $\hat{\mathbf{b}}'$ that would be obtained as above by the QR decomposition of X .

X has block-angular structure^{5,6}, advantage of which can be taken for a numerically efficient solution. The size of typical problems in the gas standards area is such that it is unnecessary to account for such structure.

^aThe solution is straightforwardly implemented in a system providing appropriate linear algebra support. For instance, in MATLAB¹¹, $[Q, R] = \text{qr}(X)$, $\text{bhatprime} = R \setminus (Q' * \mathbf{y})$.

In other areas where sets of coupled calibration curves are to be used and there are many measurements to be processed, savings in computer time and memory would be possible.

4. Uncertainties

The uncertainties associated with the use of uncorrected and corrected measurements in obtaining the calibration curves, and with the use of the calibration curves so obtained, are considered.

4.1. Uncorrected measurements

For the uncorrected measurements, the uncertainty matrix (covariance matrix) associated with $\hat{\mathbf{b}}$ is $V(\hat{\mathbf{b}}) = \sigma^2 U U^T$, where U , an upper-triangular matrix, is the inverse of R (formed by solving the triangular system $RU = I$, with I the identity matrix). The value of σ would either be provided *a priori* as the measurement standard uncertainty σ_ϵ associated with the elements of \mathbf{y} , or (if there is no lack of conformity of the model and the data) estimated by the root-mean-square residual $(S/\nu)^{1/2} = \|\mathbf{y} - X\hat{\mathbf{b}}\|_2/\nu^{1/2}$, where S is given by expression (6) and $\nu = m - n$ is the degrees of freedom.

4.2. Corrected measurements

For the corrected measurements, the uncertainty matrix $V(\hat{\mathbf{b}}')$ would be formed as above, where σ_ϵ denotes the standard uncertainty associated with $c_i y_{ij}$, i.e., the corrected y_{ij} . For a valid fit to the data, σ_ϵ can be estimated by $(S_{\text{corrected}}/(m - n + 1))^{1/2}$, where $S_{\text{corrected}}$ is given by expression (9). The degrees of freedom is $m - n + 1$, rather than $m - n$, because of the use of the resolving constraint (10).

4.3. Use of the calibration curves

The standard uncertainties associated with the values x_j^* determined from the inverse use of the calibration curve, i.e., through expression (11) are evaluated by the straightforward application of the law of propagation of uncertainty^{2,7}:

$$u^2(x_j^*) = \frac{1}{b^2} \left\{ u^2(y_j^*) + u^2(\hat{\alpha}_j) + x^2 u^2(\hat{\beta}_j) + 2x^2 u(\hat{\alpha}_j, \hat{\beta}_j) \right\}, \quad (13)$$

where $u(\hat{\alpha}_j, \hat{\beta}_j)$ is the covariance associated with $\hat{\alpha}_j$ and $\hat{\beta}_j$. $u(y_j^*)$ is taken as σ_ϵ , and $u^2(\hat{\alpha}_j)$, $u^2(\hat{\beta}_j)$ and $u(\hat{\alpha}_j, \hat{\beta}_j)$ are taken as the appropriate elements of the uncertainty matrix $V(\hat{\mathbf{b}})$ or $V(\hat{\mathbf{b}}')$.

The main contribution to $u(x_j^*)$ is typically from the term involving $u(y_j^*)$, especially if many standards are used in the calibration (i.e., N is large). In this case, $u(x_j^*) \approx u(y_j^*)/|b|$. When working with corrected measurements, additional terms should strictly be included in expression (13) that result from correlations between the curve parameters and the correction factors. The corresponding covariances are available as elements of the uncertainty matrix $V(\hat{\mathbf{b}})$; these too are expected to be small.

5. Results

Results are given for two applications of the method, to the measurement of five and then seven minor gas components.

5.1. *Five minor gas components (26 standards)*

The numerical method described in section 3 was applied to measurements of $q = 5$ minor gas components using $N = 26$ standards. Tests for normality of the residuals corresponding to uncorrected and corrected measurements (cf. section 2) were carried out to check the validity of the distributional assumption made in section 2.2. The Kolmogorov-Smirnov test⁸ was used for this purpose, under the null hypothesis that the underlying distribution is normal. The tests provided no reason to reject the hypothesis.

Tables 1, 2 and 3 show the results obtained. Table 1 gives the residual sums of squares for straight-line calibration curves for the main components of the standards without and with correction. Some highly significant improvements are observed. In particular, the residual sums of squares for the first components, nitrogen and methane (table 1), reduced by factors of 3×10^3 and 1×10^6 , respectively. The corresponding root-mean-square residuals reduced by factors of 5×10^1 and 1×10^3 . The standard uncertainties for the calculated molar fractions (table 2) reduced by the same factors (to the one significant digit quoted), as might be expected in the context of regression.

Table 3 gives the 26 estimated correction factors. They are all of the order of unity as expected. Expressed as fractional changes in the response values, they range from -8% to $+6\%$. In this example, the range of sampling conditions was artificially enlarged to demonstrate the method.

Figures 6 and 7 depict the results obtained for the first two of these components, nitrogen and methane. In each figure the uncorrected data and the (independently constructed) calibration curves are shown to the left, and the corrected data and the coupled calibration curves to the right.

Table 1. Residual sums of squares for straight-line calibration curves with and without correction for measurements of $q = 5$ gas components using $N = 26$ standards.

Gas components	Residual sum of squares	
	Experimental results	Corrected model
Nitrogen	1.8×10^9	6.6×10^5
Methane	9.7×10^{10}	8.2×10^4
CO ₂	9.8×10^8	1.1×10^6
Ethane	7.3×10^8	8.7×10^5
Propane	2.0×10^8	7.0×10^5
Total RSS	1.0×10^{11}	3.4×10^6

Table 2. Inverse use of the calibration curves obtained with and without correction factors.

Results	Gas	Responses	Calculated molar fraction x_j^*	Standard uncertainty $u(x_j^*)$
Uncorrected	Nitrogen	5.1298×10^5	19.2706	0.3539
Corrected			20.2408	0.0072
Uncorrected	Methane	1.8753×10^6	91.3191	3.3925
Corrected			90.7668	0.0028

Table 3. The correction factors corresponding to the results in table 1.

1.049	1.056	1.046	0.966	1.044	0.976	1.022	0.958	0.958
0.979	1.036	0.977	0.976	0.915	0.968	0.968	0.973	1.021
1.030	0.947	0.990	1.020	1.030	1.026	1.014	1.041	

5.2. Seven minor gas components (26 standards)

Table 4 gives a further instance of the residual sums of squares for straight-line calibration curves without and with correction. The measurement results corresponded to seven minor gas components and 26 standards. Again the Kolmogorov-Smirnov test gave no reason to doubt that the underlying distribution is normal. Although a substantial improvement was again observed, it was not as large as for the first example, implying that the systematic effects were not so dominant at the lower concentrations involved in this case.

6. Extensions

Two extensions of the approach are briefly indicated, one to higher-order models and the other to different uncertainty structures.

6.1. Higher-order models

The approach extends to quadratic (or higher-order) regression models (that are linear in the parameters) that account for possibly non-linear

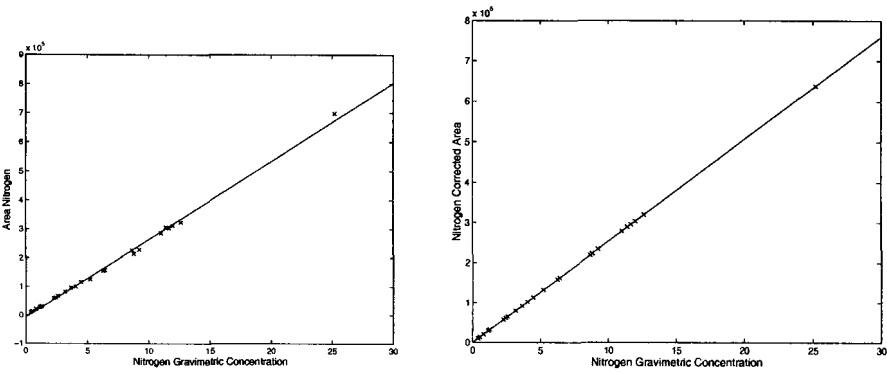


Figure 6. Experimental (uncorrected) data and calibration curve for nitrogen and the corresponding corrected results (right).

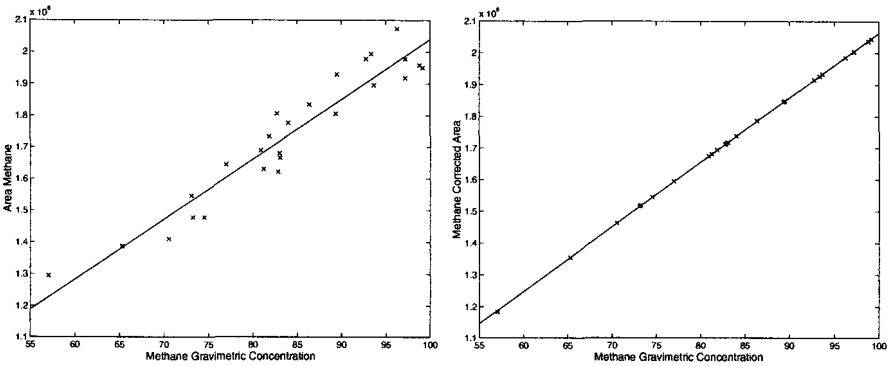


Figure 7. As figure 6 but for methane.

Table 4. The residual sums of squares for straight-line calibration curves with and without correction for measurements of $q = 7$ gas components using $N = 26$ standards.

Gas components	Residual sum of squares (RSS)	
	Experimental results	Corrected model
Propane	6.6×10^{11}	1.1×10^9
i-Butane	1.1×10^{10}	2.3×10^9
n-Butane	1.7×10^{10}	2.2×10^9
neo-Pentane	4.3×10^8	1.2×10^8
i-Pentane	8.1×10^8	4.2×10^8
n-Pentane	6.2×10^8	5.7×10^8
n-Hexane	7.3×10^8	6.2×10^8
Total RSS	7.0×10^{11}	7.3×10^9

behaviour of the detector:

$$c_i y_{ij} = \widehat{\alpha}_j + \widehat{\beta}_j x_{ij} + \widehat{\gamma}_j x_{ij}^2 + \xi_{ij}, \quad \xi_{ij} \sim NID(0, \sigma_\xi^2).$$

When the resulting calibration curves are used to obtain estimates of unknown molar fractions, it is generally necessary to use some form of ‘inverse interpolation’. Approaches for such computations and the evaluation of the associated uncertainties are available ⁷.

6.2. Other uncertainty structures

If the measurement data has different uncertainties associated with the measurement results, suitable weights can be introduced to accommodate the influence on the model parameter estimates. Define the weight matrix

$$W = \text{diag}(\mathbf{w}_1^T, \dots, \mathbf{w}_q^T), \quad \mathbf{w}_j = (w_{1j}, \dots, w_{Nj})^T.$$

In the absence of systematic effects the regression model (5) becomes

$$w_{ij}^{1/2} y_{ij} = w_{ij}^{1/2} (\widehat{\alpha}_j + \widehat{\beta}_j x_{ij}) + \varepsilon_{ij},$$

with $w_{ij} = 1/u^2(y_{ij})$, giving

$$WX\mathbf{b} \simeq W\mathbf{y}$$

as the weighted counterpart of the unweighted system $X\mathbf{b} \simeq \mathbf{y}$. The solution can be effected using QR decomposition as before, except that the matrix to be factorized is WX rather than X and the ‘right-hand side’ is $W\mathbf{y}$ rather than \mathbf{y} .

If there are systematic effects also, the regression model becomes

$$w_{ij}^{1/2} \widehat{c}_i y_{ij} = w_{ij}^{1/2} (\widehat{\alpha}_j + \widehat{\beta}_j x_{ij}) + \varepsilon_{ij},$$

with $w_{ij} = 1/(\widehat{c}_i^2 u^2(y_{ij}))$, giving

$$\frac{y_i}{u(y_i)} = \frac{\widehat{\alpha}_j}{\widehat{c}_i} \frac{1}{u(y_i)} + \frac{\widehat{\beta}_j}{\widehat{c}_i} \frac{x_{ij}}{u(y_i)} + \varepsilon_{ij}, \quad (14)$$

resulting in a minimization problem that is *non-linear* in the parameters. This problem would be solved using an appropriate non-linear least-squares algorithm ⁹ in conjunction with the resolving constraint (10). An *approximate* linear model would be given by multiplying all terms in the model (14) by c_i and, since c_i is expected to be close to unity, replacing $c_i \varepsilon_{ij}$ by ε_{ij} , yielding a counterpart of the approach for the homoscedastic case.

7. Conclusions

In a class of problems arising in the analysis of natural gases, the derived calibration curves that relate stimulus concentrations to indicated detector responses are mutually dependent as a consequence of changing environmental and instrumental conditions during measurement. A model was constructed that accounts for the systematic effects that give rise to this covariation. The method implemented to estimate the model parameters is least squares. It provides a solution that essentially removes these effects. The result was a set of corrected measurements whose associated uncertainties are dominantly associated with random effects and are significantly smaller in magnitude. The required molar fractions can then be determined appreciably more accurately compared with their determination without making such a correction. The numerical methods used to provide such solutions were outlined.

References

1. A. C. Atkinson. *Plots, Transformations and Regression*. Clarendon Press, Oxford, 1985.
2. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML. Guide to the Expression of Uncertainty in Measurement, 1995. ISBN 92-67-10188-9, Second Edition.
3. G. C. Casella and R. L. Berger. *Statistical Inference*. Duxbury Press, 2001. Second Edition.
4. D. R. Cox and D. V. Hinkley. A note on the efficiency of least-squares estimates. *J. R. Statist. Soc.*, B30:841, 1968.
5. M. G. Cox. The least-squares solution of linear equations with block-angular observation matrix. In M. G. Cox and S. J. Hammarling, editors, *Reliable Numerical Computation*, pages 227–240, Oxford, 1990. Oxford University Press.
6. M. G. Cox, A. B. Forbes, P. M. Fossati, P. M. Harris, and I. M. Smith. Techniques for the efficient solution of large scale calibration problems. Technical Report CMSC 25/03, National Physical Laboratory, Teddington, UK, 2003.
7. M. G. Cox, A. B. Forbes, P. M. Harris, and I. M. Smith. The classification and solution of regression problems for calibration. Technical Report CMSC 24/03, National Physical Laboratory, Teddington, UK, 2003.
8. M. H. DeGroot. *Probability and Statistics, 3rd Ed.* Addison-Wesley, Reading, Massachusetts, 1991.
9. P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, London, 1981.
10. N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, 1996.
11. The MathWorks. *Matlab Users Manual*. MathWorks Inc., Natick, Mass., USA, 1992.

PARAMETRIZED APPROXIMATION ESTIMATORS FOR MIXED NOISE DISTRIBUTIONS*

D. P. JENKINSON, J. C. MASON, A. CRAMPTON
M. G. COX*, A. B. FORBES*, R. BOUDJEMAA*

School of Computing and Engineering, University of Huddersfield, UK

**Centre for Mathematics and Scientific Computing,*

National Physical Laboratory, Teddington, UK

Email: {d.p.jenkinson, j.c.mason, a.crampton}@hud.ac.uk

{maurice.cox, alistair.forbes, redouane.boudjemaa}@npl.co.uk

Consider approximating a set of discretely defined values f_1, f_2, \dots, f_m say at $x = x_1, x_2, \dots, x_m$, with a chosen approximating form. Given prior knowledge that noise is present and that some might be outliers, a standard least squares approach based on an l_2 norm of the approximation error ϵ may well provide poor estimates. We instead consider a least squares approach based on a modified measure taking the form $\tilde{\epsilon} = \epsilon(1 + c^2\epsilon^2)^{-1/2}$, where c is a constant to be fixed. Given a prior estimate of the likely standard deviation of the noise in the data, it is possible to determine a value of c such that the estimator behaves like a robust estimator when outliers are present but like a least squares estimator otherwise. We describe algorithms for computing the parameter estimates based on an iteratively weighted linear least squares scheme, the Gauss-Newton algorithm for nonlinear least squares problems and the Newton algorithm for function minimization. We illustrate their behaviour on approximation with polynomial and radial basis functions and in an application in co-ordinate metrology.

1. Introduction

Approximation is concerned with fitting a function $F(x, \mathbf{a})$ depending on parameters \mathbf{a} to a set of data (x_i, f_i) , $i = 1, \dots, m$, in such a way as to minimize over \mathbf{a} some measure of the error of fit. The form F is often specified as a linear combination

$$F(x, \mathbf{a}) = \sum_{j=1}^n a_j \phi_j(x),$$

*Work partially funded under EU SofTools_MetroNet Contract N. G6RT-CT-2001-05061

of basis functions $\phi_j(x)$. For example, we can consider a polynomial basis where $\phi_j(x) = x^{j-1}$, or $\phi_j(x) = T_{j-1}(x)$ (Chebyshev basis) or a radial basis function (RBF) with $\phi_j(\mathbf{x}) = \phi(\|\mathbf{x} - \boldsymbol{\lambda}_j\|)$, where ϕ is a univariate function and $\|\bullet\|$ denotes the Euclidean norm. (Thus, $\|\mathbf{x} - \boldsymbol{\lambda}_j\|$ is simply the distance from \mathbf{x} to the centre $\boldsymbol{\lambda}_j$.)

It is quite common that noise of more than one type should be present in the data. We are particularly interested in situations where the majority of the data arising from a physical system behaves as predicted by the statistical model, but a minority (no more than 25%) has been corrupted by unknown and potentially large influence factors. The latter may not be easy to identify and we look for approximation methods that are robust with respect to such effects.

The quality of the approximation F to f is assessed by a measure, such as a norm $\|\mathbf{f} - \mathbf{F}\|$ of the approximation error $\{f_i - F(x_i, \mathbf{a})\}$. The most commonly adopted norm is the l_2 or least squares norm

$$\|\mathbf{f} - \mathbf{F}\|_2 = \left[\sum_{i=1}^m (f_i - F(x_i, \mathbf{a}))^2 \right]^{\frac{1}{2}}.$$

The l_2 norm is appropriate when the noise in the data is drawn from a normal distribution and it can be shown that a least squares approach is optimal for this case. However, if the data has wild points, i.e., large isolated errors, a least squares approach is likely to provide a poor fit. In this situation, a better norm to adopt is the l_1 norm,

$$\|\mathbf{f} - \mathbf{F}\|_1 = \sum_{i=1}^m |f_i - F(x_i, \mathbf{a})|,$$

which in general is less affected by the presence of wild points. Although l_1 approximation can be considered quite effective^{1,7}, we give three reasons for looking for other approaches. Firstly, while there exist satisfactory l_1 approximation algorithms for linear models, the algorithms for nonlinear models¹⁰ are much more difficult to implement. Secondly, l_1 approximation does not enjoy the same (or equivalent) optimal behaviour as that of l_2 approximation. The price of being able to deal with wild points is that somewhat poorer use is made of the valid data points. Thirdly, there is no satisfactory method of estimating the uncertainty associated with the fitted parameters for l_1 approximation. By comparison, the l_2 case is straightforward and well-rooted in probability theory. Ideally we aim to find a compromise between these two norms in order to provide effective algorithms for data reflecting both normally distributed noise and wild points.

In section 2 we describe a new least squares approach – *asymptotic least squares* (ALS) – that combines least squares approximation with a nonlinear transformation function that offers, we believe, a straightforward and effective method. In section 3, we present a number of algorithms for ALS approximation based on standard nonlinear least squares and iterative weighting schemes. A number of example applications are presented in sections 4–6, and our concluding remarks are given in section 7.

2. Asymptotic least squares

Below, we let (x_i, f_i) , $i = 1, \dots, m$, be a set of data points, $F(x, \mathbf{a})$ an approximating function depending on parameters $\mathbf{a} = (a_1, \dots, a_n)^T$, and $\epsilon_i = \epsilon_i(\mathbf{a}) = f_i - F(x_i, \mathbf{a})$, the approximation error associated with the i th data point. Often, F represents a linear model so that $F_i = F(x_i, \mathbf{a}) = \mathbf{c}_i^T \mathbf{a}$ for some vector of coefficients $\mathbf{c}_i = (c_{i1}, \dots, c_{in})^T$. If F is a linear combination of basis functions then $c_{ij} = \phi_j(x_i)$. In matrix notation we write $\mathbf{F} = C\mathbf{a}$ where $\mathbf{F} = (F_1, \dots, F_m)^T$.

The least-squares best fit function is the one specified by the parameters \mathbf{a} that minimize

$$E_2(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^m \epsilon_i^2(\mathbf{a}).$$

(The fraction $1/2$ is included to simplify later expressions.) The fact that the contribution to the objective function E_2 associated with each data point is the square of the approximation error means that a least squares solution will try to accommodate wild points to the detriment of the quality of fit at the remaining points. In l_1 approximation, the contribution of each point to the objective function is the approximation error itself. This allows an l_1 approximant to tolerate large approximation errors associated with wild points.

The main idea of asymptotic least squares (ALS) is to apply a transformation function to the approximation error ϵ_i that will reduce the effect of large errors associated with wild points. We look to minimize an objective function of the form

$$\tilde{E}(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^m \tilde{\epsilon}_i^2(\mathbf{a}), \quad \tilde{\epsilon}_i = \tau(\epsilon_i), \quad (1)$$

for some suitable transformation τ . We require i) τ to have continuous second derivatives so that minimizing \tilde{E} is a smooth optimization problem, ii) $\tau(0) = 0$, $\tau'(0) = 1$ and $\tau''(0) = 0$ so that for small errors, \tilde{E} has

similar behaviour to a standard least squares objective function, and iii) $\lim_{|\epsilon| \rightarrow \infty} \tau'(\epsilon) = 0$ so that increasing an already large error will have a marginal effect on \tilde{E} . A simple function satisfying these criteria is

$$\tau(x) = x/(1 + c^2 x^2)^{1/2}. \quad (2)$$

We note that $\lim_{\epsilon \rightarrow \pm\infty} \tau(\epsilon) = \pm 1/c$. There are various other choices that may be made for $\tau(\epsilon)$, with similar properties to (2), e.g., $\tau(\epsilon) = \tanh(c\epsilon)$.

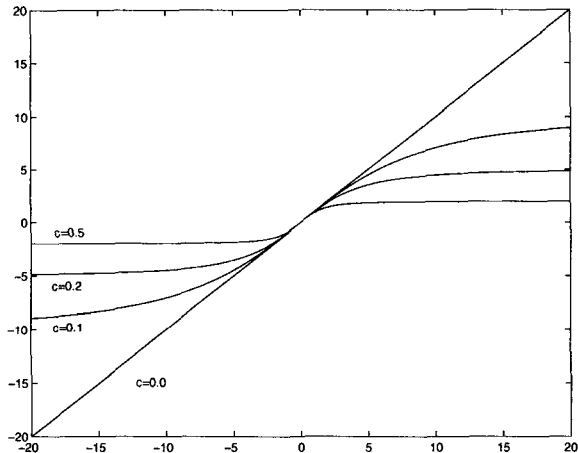


Figure 1. Graph of τ defined in (2) for different values of c .

The parameter c in (2) controls the level of ϵ at which the transform takes effect (figure 1). If we expect the (standard) noise in the data to be described by a normal distribution with mean zero and standard deviation unity, then we would expect approximately 95% of the approximation errors to lie in the interval $[-2, 2]$. In this region, we want τ to make a small change, suggesting a value of c in the region of $c = 1/4$.

Given such a τ satisfying the criteria above, we can define related transform functions of the form

$$\tau(x) = \begin{cases} x, & |x| \leq d, \\ \tau(x - d) + d, & x > d, \\ \tau(x + d) - d, & x < -d, \end{cases} \quad (3)$$

that also satisfy the criteria. The parameter d ensures that $\tilde{\epsilon}$ is the same as l_2 for errors in $[-d, d]$.

The definition of τ in (3) can be compared with a Huber estimation approach⁶ which minimizes

$$E_H(\mathbf{a}) = \sum_{i=1}^m \rho(\epsilon_i(\mathbf{a})),$$

where

$$\rho(x) = \begin{cases} x^2, & \text{if } |x| \leq d, \\ d|x|, & \text{if } |x| > d. \end{cases}$$

For small residuals (data generated by a system obeying the statistical model), the Huber estimator behaves like a least squares estimator. For large residuals (outliers) the estimator is like ℓ_1 .

If the noise in the data is drawn from a distribution with standard deviation σ then an appropriate form of τ is

$$\tau(x) = x/(1 + c^2 x^2/\sigma^2)^{1/2} \quad \text{or} \quad \tau(x) = (x/\sigma)/(1 + c^2(x/\sigma)^2)^{1/2}. \quad (4)$$

In terms of determining the parameters \mathbf{a} these two forms are equivalent but the latter formulation gives a more statistically relevant form.

3. Algorithms for asymptotic least squares

Even if F is linear in the parameters \mathbf{a} the introduction of the nonlinear τ function makes the minimization of \tilde{E} a nonlinear least squares problem. In this section we consider general approaches to its solution.

3.1. Nonlinear least squares optimization

Let us first consider the problem of minimizing a general (smooth) function $E(\mathbf{a})$ of n variables⁴. Let $\mathbf{g}(\mathbf{a})$ be the gradient of E , $g_j = \partial E/\partial a_j$, and H the Hessian matrix of second partial derivatives, $H_{jk} = \partial^2 E/\partial a_j \partial a_k$. In the Newton algorithm, an estimate of the solution \mathbf{a} is updated according to $\mathbf{a} := \mathbf{a} + t\mathbf{p}$, where \mathbf{p} solves $H\mathbf{p} = -\mathbf{g}$ and t is a step length chosen to ensure a sufficient decrease in E . Near the solution, the Newton algorithm converges quadratically. Away from the solution, there is no guarantee that the Hessian matrix will be strictly positive definite and algorithms have to deal appropriately with this eventuality⁴.

Suppose now E is a sum of squares function $E(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^m e_i^2(\mathbf{a})$ and let $J(\mathbf{a})$ be the Jacobian matrix $J = \partial e_i/\partial a_j$. Then $\mathbf{g} = J^T \mathbf{e}$ and $H = J^T J + G$, where

$$G_{jk} = \sum_{i=1}^m e_i \frac{\partial^2 e_i}{\partial a_j \partial a_k}.$$

The Gauss-Newton (GN) algorithm follows the same approach as the Newton algorithm, only that in determining the update step, H is approximated by $J^T J$, i.e, the term G is ignored and \mathbf{p} is found by solving the linear least squares problem $J^T J \mathbf{p} = -J^T \mathbf{e}$. A stable method of solving this system^{3,5} is to use an orthogonal factorization of the matrix J . The GN algorithm in general converges linearly at a rate that depends on the condition of the approximation problem, the size of the residual errors near the solution and the curvature. If the problem is well-conditioned, the residuals are small and the summand functions e_i are nearly linear, then $J^T J$ is a good approximation to the Hessian matrix H and convergence is fast. The GN algorithm has the advantage that if J is full rank, then $J^T J$ is strictly positive definite so that some of the complications associated with a Newton algorithm implementation are avoided.

Uncertainty matrix associated with the fitted parameters. For standard nonlinear least squares problems, the uncertainty matrix of the fitted parameters³ is estimated by

$$U_{\mathbf{a}} = \hat{\sigma}^2 (J^T J)^{-1},$$

where $\hat{\sigma}$ is an estimate of the standard deviation σ of the noise in the data. A posterior estimate is $\hat{\sigma}^2 = \mathbf{e}^T \mathbf{e} / (m - n)$, where \mathbf{e} the vector of residuals evaluated at the solution. If the Hessian matrix H is available, $U_{\mathbf{a}} = \hat{\sigma}^2 H^{-1}$ is generally a better estimate.

Let us now return to the ALS problem of minimizing $\tilde{E}(\mathbf{a})$ defined in (1). To employ a Newton-type algorithm we need to calculate

$$\mathbf{g} = \tilde{J}^T \tilde{\epsilon}, \quad \tilde{J}_{ij} = \dot{\tau}_i \frac{\partial \epsilon_i}{\partial a_j}, \quad \dot{\tau}_i = \frac{d\tau}{dx}(\epsilon_i),$$

and

$$\tilde{H} = \tilde{J}^T \tilde{J} + \tilde{G}, \quad \tilde{G}_{jk} = \sum_i \tilde{\epsilon}_i \frac{\partial^2 \tilde{\epsilon}_i}{\partial a_j \partial a_k}.$$

We note that

$$\frac{\partial^2 \tilde{\epsilon}_i}{\partial a_j \partial a_k} = \ddot{\tau}_i \frac{\partial \epsilon_i}{\partial a_j} \frac{\partial \epsilon_i}{\partial a_k} + \dot{\tau}_i \frac{\partial^2 \epsilon_i}{\partial a_j \partial a_k}, \quad \ddot{\tau}_i = \frac{d^2 \tau}{dx^2}(\epsilon_i).$$

The first term on the right is the contribution due to the curvature in τ , the second, due to that in F . Even if the second term is small, the first term is likely to be significant. This means that in practice the Gauss-Newton algorithm implemented for ALS will have significantly slower convergence than a Newton algorithm. However, if F is linear with $\mathbf{F} = \mathbf{f} - \mathbf{C}\mathbf{a}$, the

second term is zero and a Newton algorithm can be implemented easily with \tilde{J} and \tilde{G} calculated using the following identities:

$$\tilde{J}_{ij} = -c_{ij}\tilde{\tau}_i, \quad \tilde{G}_{jk} = \sum_i \tau_i \tilde{\tau}_i c_{ij} c_{ik}.$$

3.2. Iteratively weighted least squares

In this section, we introduce two novel algorithms for tackling the nonlinear problem of minimizing $\tilde{E}(\mathbf{a})$. Iterating over q , we minimize at step q ,

$$\sum_{i=1}^m \left(w_i^q \epsilon_i^{(q+1)}(\mathbf{a}) \right)^2, \quad w_i^q = \left(1 + (c\epsilon_i^{(q)})^2 \right)^{-1/2}, \quad q = 0, 1, \dots, \quad (5)$$

to determine estimates \mathbf{a}_{q+1} and associated approximation errors $\epsilon_i^{(q+1)}$. Here, the weights w_i^q are fixed and determined from the solution $\epsilon_i^{(q)}$ at the q th iteration. Updated estimates of the parameters \mathbf{a} are therefore found by solving a weighted least squares problem. In practice this algorithm converges at a rapid rate to a near-best l_2 approximation. Since each step is a least squares problem the algorithm is straightforward to implement, especially for linear models.

Having observed the simple but general form of (5), we are able, as an alternative, to replace this “multiplicative” iteration with a similar, but additive, iteration:

$$\sum \left[\epsilon_i^{(q+1)}(\mathbf{a}) + \tilde{\epsilon}_i^{(q)} - \epsilon_i^{(q)} \right]^2. \quad (6)$$

This is very similar to an algorithm proposed by Mason and Upton⁹. If we write $\delta_i^q = \left(1 + (c\epsilon_i^{(q)})^2 \right)^{1/2} = 1/w_i^q$, the form simplifies to

$$\sum_{i=1}^m \left[\epsilon_i^{(q+1)}(\mathbf{a}) - (c\epsilon_i^{(q)})^2 \left\{ \frac{\epsilon_i^{(q)}}{\delta_i^q + (\delta_i^q)^2} \right\} \right]^2.$$

Again it is observed that the algorithm converges rapidly in practice to a near-best approximation.

3.3. Maximum likelihood approaches

The effectiveness of the ALS approach depends on having a suitable estimate of the standard noise in the data. In this section we describe an approach borrowed from maximum likelihood estimation^{2,8} to provide such

an estimate. We consider the function $\tau(x) = (x/\sigma)/(1 + c^2(x/\sigma)^2)^{1/2}$, where now c is fixed but σ has to be determined. Writing

$$\tilde{\epsilon}_i = w_i \epsilon_i, \quad w_i = w_i(\sigma) = \frac{1}{(\sigma^2 + c^2 \epsilon_i^2)^{1/2}},$$

we can regard the ALS as a nonlinearly weighted least squares approach where the weights are parametrized by σ . We determine estimates of σ and \mathbf{a} by solving the optimization problem

$$\min_{\mathbf{a}, \sigma} \left\{ \frac{1}{2} \sum_{i=1}^m \tilde{\epsilon}_i^2(\mathbf{a}, \sigma) - \sum_{i=1}^m \log w_i(\sigma) \right\}. \quad (7)$$

As in maximum likelihood estimation, the solution estimates of \mathbf{a} and σ are those that are the most likely to give rise to the data $\{(x_i, f_i)\}$. If σ is known, then the objective function in (7) simplifies to a least squares objective function (1).

4. Example: polynomial approximation

In this section, we discuss the behaviour of ALS methods on polynomial regression. The model function is linear and of the form

$$F(x, \mathbf{a}) = \sum_{j=1}^n a_j \phi_j(x),$$

where ϕ_j is a Chebyshev polynomial basis function of degree $j-1$. Figure 2 gives an example of least squares (LS) and asymptotic least squares (ALS) fits to 26 data points with 5 outliers (i.e., approximately 20% outliers). The figure shows that the ALS fit follows the majority of the data while the LS fit is skewed by the wild points.

We use Monte Carlo simulation to compare ALS with LS approximation. We generate sets of data as follows. We fix m the number of data points, $m_O < m$ the number of wild points, n the number of basis functions, σ the standard deviation for standard noise and $\sigma_O > \sigma$ the standard deviation associated with wild points. We form uniformly spaced abscissae $\mathbf{x} = (x_1, \dots, x_m)^T$ and corresponding observation matrix C with $c_{ij} = \phi_j(x_i)$. Given $\mathbf{a} = (a_1, \dots, a_n)^T$ and number N of Monte Carlo simulations, we perform the follow steps.

- I Set $\mathbf{f}_0 = C\mathbf{a}$. The points $(x_i, f_{i,0})$ lie on the polynomial curve specified by parameters \mathbf{a} .
- II For each $q = 1, \dots, N$,

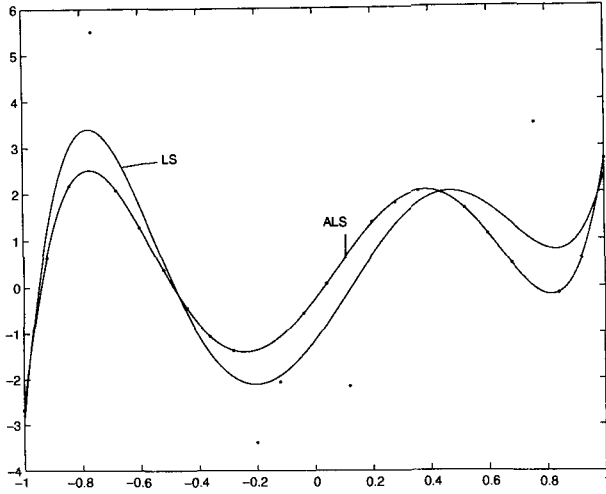


Figure 2. Example LS and ALS polynomial fits to data with outliers. The ALS fit follows the majority of the data while the LS fit is skewed by the wild points.

- i Add standard noise: $\mathbf{f}_q = \mathbf{f}_0 + \mathbf{r}$, $r_i \in N(0, \sigma^2)$.
- ii Choose at random a subset $I_q \subset \{1, \dots, m\}$ of m_O indices.
- iii Add large perturbations: for each $i \in I_q$, $\mathbf{f}_q(i) = \mathbf{f}_q(i) + r_{i,O}$, $r_{i,O} \in N(0, \sigma_O^2)$.
- iv Find the least squares and asymptotic least squares fits to data set $\{(x_i, f_{i,q})\}$ determining parameter estimates \mathbf{a}_q and $\tilde{\mathbf{a}}_q$, respectively.
- v Calculate estimate of the standard deviation of the residuals:

$$\hat{\sigma}_q^2 = \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} / (m - n), \quad \tilde{\sigma}_q^2 = \tilde{\boldsymbol{\epsilon}}^T \tilde{\boldsymbol{\epsilon}} / (m - n),$$

$$\text{with } \boldsymbol{\epsilon} = \mathbf{f}_q - C\mathbf{a}_q, \quad \tilde{\boldsymbol{\epsilon}} = \tau(\mathbf{f}_q - C\tilde{\mathbf{a}}_q).$$

- vi Store estimates: $A(q, 1 : n) = \mathbf{a}_q^T$, $\tilde{A}(q, 1 : n) = \tilde{\mathbf{a}}_q^T$, $S(q) = \hat{\sigma}_q$ and $\tilde{S} = \tilde{\sigma}_q$.

From these simulations we can compare, for example, the sample standard deviations $u_{j,LS}^{MC}$ and $u_{j,ALS}^{MC}$ of the calculated parameter values $\{A(i, j)\}_{i=1}^m$ and $\{\tilde{A}(i, j)\}_{i=1}^m$, respectively, with that estimated using the standard approach described in 3.1.

Table 1 gives the uncertainties for the case $m = 51$, $m_O = 5$, $n = 7$, $\sigma = 0.0005$, $\sigma_O = 0.01$ and $N = 5000$. The ALS fits were determined using $\tau(x) = x(1 + c^2 x^2)^{-1/2}$ with $c = (4\sigma)^{-1}$. The table presents the standard uncertainties associated with the fitted parameters calculated in a number

of ways: u_{LS} – LS fit using $\hat{\sigma}$, u_{LS}^0 – LS using σ , u_{LS}^{MC} – LS Monte Carlo estimate, \bar{u}_{ALS} – ALS using $\tilde{\sigma}$ and approximate Hessian matrix, u_{ALS} – ALS using $\tilde{\sigma}$ and exact Hessian matrix, u_{ALS}^0 – ALS using σ and exact Hessian matrix, u_{ALS}^{MC} – ALS Monte Carlo estimate. The table indicates:

- i) For the LS fit, the estimates u_{LS} using the posterior estimate $\hat{\sigma}$ are in good agreement with the Monte Carlo estimates u_{LS}^{MC} . This reflects the fact that for linear problems, the posterior estimate of σ is reasonably reliable.
- ii) The estimates u_{LS}^0 using the prior estimate σ significantly underestimate the variation. These estimates would be valid if there were no outliers.
- iii) For the ALS fit, there is a modest difference between the estimates \bar{u}_{ALS} using the approximate Hessian $\tilde{J}^T \tilde{J}$ and those u_{ALS} using the exact Hessian $\tilde{H} = \tilde{J}^T \tilde{J} + \tilde{G}$.
- iv) The estimates using the posterior estimate $\tilde{\sigma}$ are larger than the Monte Carlo estimate u_{ALS}^{MC} while the estimate u_{ALS}^0 using the prior estimate σ is in reasonable agreement with the Monte Carlo estimates. The average value for $\tilde{\sigma}$ is approximately 0.00074 compared with $\sigma = 0.0005$.
- v) The Monte Carlo estimates u_{ALS}^{MC} are only modestly larger than u_{LS}^0 , the uncertainties estimates we would expect if there were no outliers present. They are approximately a factor of five smaller than those for the LS fits.

The results show firstly the effectiveness of ALS approximation relative to LS and secondly, the uncertainty estimates associated with the ALS fits are a reasonable guide.

These results have been generated using a value of c based on the known standard deviation σ associated with the standard noise. In practice, we may only have a very approximate estimate of σ . We have also applied a maximum likelihood (ML) approach (section 3.3) to this type of data, using the LS estimates of \mathbf{a} and posterior estimate $\hat{\sigma}$ as starting estimates. The average ML estimate of σ is approximately 0.00057 and compares well with the assigned value of 0.0005. The ML estimates of \mathbf{a} are close to those determined by ALS with the assigned value of σ . While the ML approach requires general optimization algorithms rather than least squares algorithms, it may well be valuable in situations where little is known about the underlying noise.

Table 1. Estimates of the standard uncertainties of the fitted parameters: u_{LS} – LS fit using $\hat{\sigma}$, u_{LS}^0 – LS using σ_0 , u_{LS}^{MC} – LS Monte Carlo estimate, \bar{u}_{ALS} – ALS using $\hat{\sigma}$ and approximate Hessian matrix, u_{ALS} – ALS using $\hat{\sigma}$ and exact Hessian matrix, u_{ALS}^0 – ALS using σ and exact Hessian matrix, u_{ALS}^{MC} – ALS Monte Carlo estimate.

	u_{LS}	u_{LS}^0	u_{LS}^{MC}	\bar{u}_{ALS}	u_{ALS}	u_{ALS}^0	u_{ALS}^{MC}
a_1	1.0219	0.1616	1.0400	0.2616	0.2821	0.1907	0.1936
a_2	0.8474	0.1340	0.8532	0.2137	0.2316	0.1565	0.1629
a_3	0.7837	0.1239	0.7925	0.1976	0.2126	0.1437	0.1554
a_4	0.7378	0.1167	0.7545	0.1904	0.2046	0.1383	0.1439
a_5	0.7118	0.1126	0.7167	0.1766	0.1845	0.1247	0.1354
a_6	0.6615	0.1046	0.6607	0.1699	0.1800	0.1217	0.1208
a_7	0.6508	0.1029	0.6489	0.1668	0.1774	0.1199	0.1207

5. Example: approximation with radial basis functions

In this example we fit a radial basis surface to data. The radial basis function (RBF) has centres λ_j lying on a regular 5×5 grid, a subgrid of a 21×21 grid of points \mathbf{x}_i . The observation matrix was defined as

$$C_{ij} = \exp\{-\|\mathbf{x}_i - \lambda_j\|^2 / (2\rho^2)\},$$

associated with a Gaussian RBF. Given \mathbf{a} , a vector of coefficients, the height \mathbf{f}_0 of the surface at the grid points was calculated from $\mathbf{f}_0 = \mathbf{C}\mathbf{a}$. Figure 3 graphs the resulting surface. Random perturbations with standard deviation $\sigma = 0.001$ were added to \mathbf{f}_0 to generate heights \mathbf{f} with $f_i = f_{i,0} + r_i$, $r_i \in N(0, \sigma^2)$. To simulate wild points, we set $\mathbf{f}_O = \mathbf{f}$ and at 40 random locations $i \in I_O \subset \{1, \dots, 441\}$, large perturbations were added $f_{i,O} = f_i + r_{i,O}$, $r_{i,O} \in N(0, \sigma_O^2)$, with $\sigma_O = 0.020$. In Figure 4 we graph the difference between the LS fit to the heights \mathbf{f}_O (with wild points) with that to heights \mathbf{f} (with standard noise). Figure 5 plots the same information for the ALS fit. For the LS fit the wild points skew the fitted surface by up to 0.006 (much larger than σ) while for the ALS fit the difference is smaller than 0.0004 (less than $\sigma/2$). The ALS approach is much better at approximating the underlying surface in the presence of wild points.

We use this example to compare the convergence of three optimization algorithms, the iteratively reweighted algorithm (IW) discussed in section 3.2 based on equation (5) and the Gauss-Newton (GN) and Newton (N) algorithms discussed in section 3.1. Table 2 shows the norm of the step taken (change in the vector of parameter estimates) at each iteration. Both the IW and GN algorithms show linear convergence behaviour with similar rates of convergence. The Newton algorithm exhibits quadratic convergence.

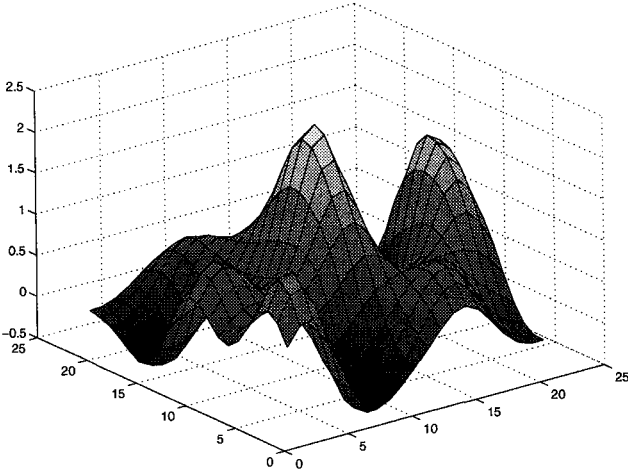


Figure 3. Surface generated using a Gaussian radial basis function.

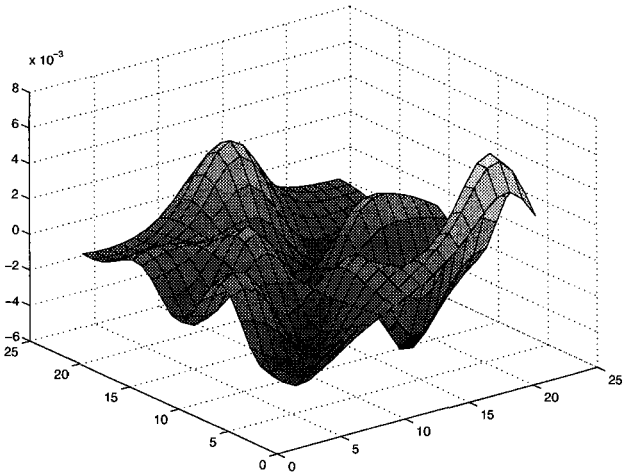


Figure 4. Difference between the LS RBF approximation to data with wild points (mixed noise) and data with standard noise.

6. Example: assessment of aspheric surfaces

In determining the shape of high quality optical surfaces using measurements gathered by a coordinate measuring machine, care must be taken

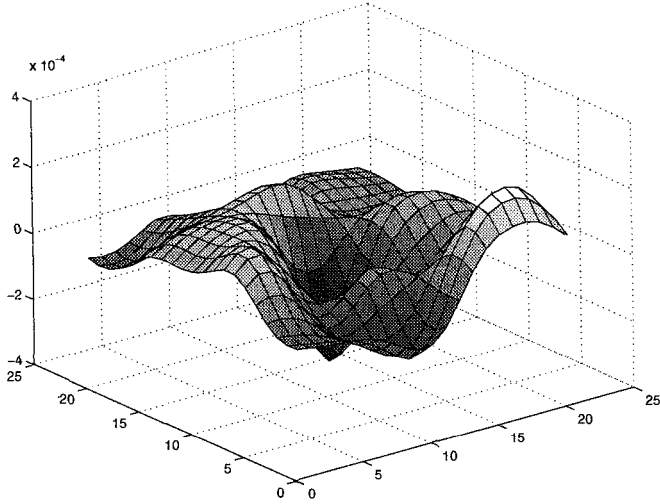


Figure 5. Difference between the ALS RBF approximation to data with wild points (mixed noise) and data with standard noise.

Table 2. Convergence behaviour of the IW, GN and N algorithms for approximation with radial basis functions to data with wild points. For each algorithm the norm $\|\mathbf{p}_q\| = \|\mathbf{a}_q - \mathbf{a}_{q_1}\|$ of the change in parameter values at the q th iteration. Each algorithm starts with estimates \mathbf{a}_0 , the solution to the standard least squares problem.

q	IW	GN	N
1	5.8686e-003	9.9268e-003	5.9427e-003
2	7.6633e-004	3.5497e-003	8.5095e-004
3	1.6589e-004	1.5802e-004	3.9609e-005
4	4.1534e-005	1.6233e-005	2.8828e-008
5	1.0467e-005	2.4621e-006	1.6620e-014
6	2.6355e-006	5.0043e-007	
7	6.6184e-007	1.0022e-007	
8	1.6578e-007	1.9966e-008	
9	4.1434e-008	3.9706e-009	
10	1.0339e-008	7.8927e-010	
11	2.5770e-009		
12	6.4177e-010		

to ensure that the optical surface is not damaged by the contacting probe. However, using a low-force probing scheme, the presence of particles of dust on the artefact's surface introduces sporadic, large non-random effects into

the measurement data. Figure 6 shows the residuals associated with an ALS fit of a hyperboloid surface to measurements of an aspheric mirror, a component in an earth observation camera. The spikes are due to particles of dust on the mirror or on the spherical probe. It is judged that 9 of the 401 measurements (i.e., approximately 2%) have been contaminated. Because the dust particles must necessarily have a positive diameter an asymmetric transform function τ was used in which only large, positive approximation errors are transformed. The standard noise associated with the measurements is of the order of 0.000 2 mm while the diameter of the dust particles is of the order of 0.002 mm. The difference between the ALS fitted surface and that generated using a standard (nonlinear) approach was of the order of 0.000 4 mm, and is seen to be significant relative to the standard noise.

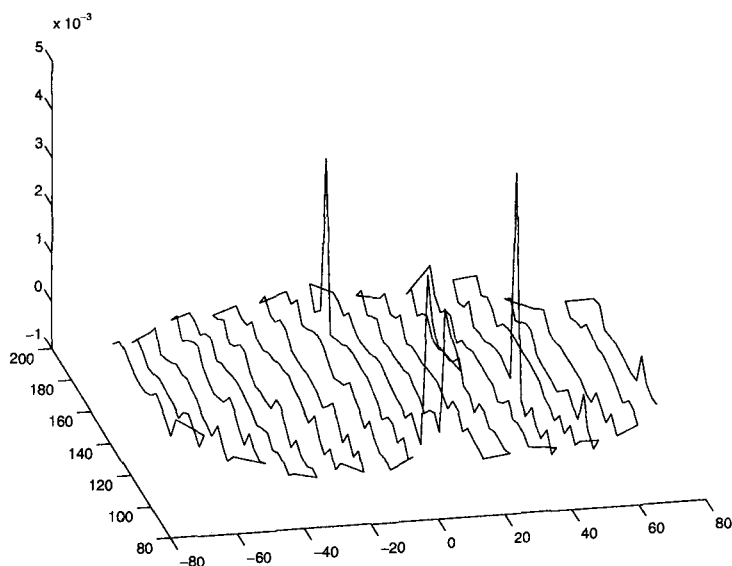


Figure 6. Residuals associated with an ALS fit of a hyperboloid surface to measurements of an aspheric mirror. The spikes are due to particles of dust on the mirror or on the spherical probe. The units for each axis are millimetres.

7. Concluding remarks

In this paper we have described a simple approach, asymptotic least squares (ALS), to approximation to data associated with mixed noise. The approach is based on least squares and uses a transform function with appro-

appropriate asymptotic behaviour. The estimates of the fitted parameters are found by solving a nonlinear least squares problem and we have described solution algorithms based on standard optimization techniques and also iteratively re-weighting approaches. These algorithms are easy to implement using standard least squares computational components.

The ALS approach can be applied to any situation in which a standard least squares analysis is appropriate. We have demonstrated its behaviour in approximation with polynomials, radial basis functions and aspheric surfaces and shown that the approach is effective in dealing with wild points. It is straightforward to generate uncertainty estimates associated with the fitted parameters but further work is required to improve and validate these approaches. It is possible that a maximum likelihood approach may prove useful.

References

1. I. Barrodale and F. D. K. Roberts. An efficient algorithm for discrete l_1 linear approximation with linear constraints. *SIAM Journal of Numerical Analysis*, 15:603 – 611, 1978.
2. M. G. Cox, A. B. Forbes, J. Flowers, and P. M. Harris. Least squares adjustment in the presence of discrepant data. *This volume*.
3. M. G. Cox, A. B. Forbes, and P. M. Harris. Software Support for Metrology Best Practice Guide 4: Modelling Discrete Data. Technical report, National Physical Laboratory, Teddington, 2000.
4. P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, London, 1981.
5. G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, third edition, 1996.
6. P. J. Huber. *Robust Statistics*. Wiley, New York, 1980.
7. D. Lei. *Robust and Efficient Algorithms for l_1 and l_∞ Problems*. PhD thesis, University of Huddersfield, 2002.
8. K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.
9. J. C. Mason and N. K. Upton. Linear algorithms for transformed linear forms. In C. K. Chui, L. L. Schumaker, and J. D. Ward, editors, *Approximation Theory VI, Volume 2*, pages 417–420, New York, 1989. Academic Press.
10. W. Murray and M. L. Overton. A projected Lagrangian algorithm for nonlinear l_1 optimization. *SIAM Journal for Scientific and Statistical Computing*, 2:207–224, 1981.

ALGORITHMS FOR THE CALIBRATION OF LASER-PLANE SENSORS ON CMMs*

C. LARTIGUE^{1,2}, P. BOURDET¹, L. MATHIEU^{1,3} AND C. MEHDI-SOUZANI¹

¹*LURPA-ENS de Cachan,
61, av Pt Wilson,
94235 Cachan cedex, France
E-mail: name@lurpa.ens-cachan.fr*

²*IUT de Cachan, 9 av de la Division Leclerc,
BP 140, 94234 Cachan cedex, France*

³*CNAM, 5 bd Descartes,
77 420 Champs sur Marne, France*

This paper deals with a method for the calibration of optical sensors on CMMs that relies on the measurement of specific artefacts. The calibration process consists of establishing the global transfer function between the 2D data given in the 2D space of the sensor and the 3D coordinates of points in the CMM space. The identification of the transfer function parameters requires the measurement of geometrical artefacts. In the paper, we suggest a specific artefact, the facet sphere. Algorithms for parameter identification are thus developed, based on the extraction of interest points belonging to the artefact. An estimation of dispersions associated with the method highlights the effect of some configuration parameters of the measuring system such as the position in the 2D space.

1. Introduction

All measuring methods using Coordinate Measuring Machines (CMMs) equipped with various sensor configurations or various sensor orientations require algorithms for the registration of point sets independently acquired.

The issue consists of identifying the transformations that allow expressing the coordinates of the different point sets obtained via different measurements in a unique coordinate system. The problem is denoted as the calibration, and concerns the choice of the identification method and the definition of associated algorithms. Most generally, the method of identification relies on the measurement of artefacts¹. In addition, the map of the measurement

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

uncertainties is established in the CMM workspace. Even though these problems are quite well controlled within the context of mechanical contact probes², they remain uncontrolled when using optical sensors, such as a laser-plane sensor. Indeed, when measuring with mechanical probes, calibration methods in 3D are based on the measurement of a reference sphere that guarantees the registration of the different measurements within the CMM coordinate system. Uncertainties can be evaluated in the whole CMM workspace, and are supposed to be independent of the sensor orientation³. When using optical sensors, existing calibration methods generally allow the calibration of the sensor itself (optical calibration)⁴.

This paper deals with a calibration method for CMMs equipped with a laser-plane sensor. For the acquisition system we consider a motorized indexing head PH10 from Renishaw (<http://www.renishaw.com>), which supports the sensor enabling the probe to be oriented to repeatable positions increasing its accessibility space. Note that this configuration is as yet seldom used. With such a measuring means, the principle of 3D point determination is triangulation. Therefore, the sensor consists of a transmitter and a receiver. The transmitter sends a thin laser beam whose width allows a large part of the surface object to be covered; the receiver, a CCD (Charged Couple Device) camera, visualizes the plane under incidence, with a given fixed triangulation angle. The CCD camera thus observes the result of the intersection between the plane laser and the object surface as a 2D image (Figure 1).

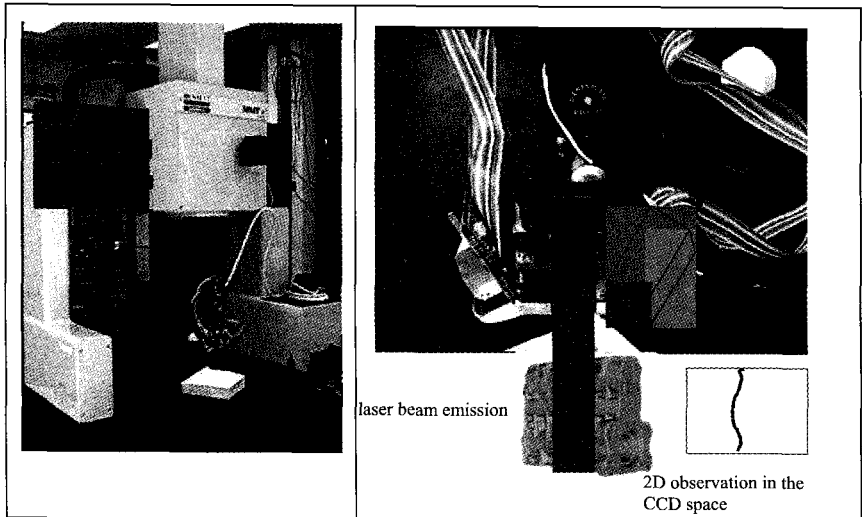


Figure 1. Measuring system and principle of acquisition.

The 3D sensor used in our developments is the system KLS51 from Kréon Technologies (<http://www.kreon3D.com>). Following the acquisition is the calibration process, which defines the relationship between the 2D data and the 3D coordinates of points belonging to the object surface. The calibration process thus conditions the accuracy of point acquisition⁵.

The paper is organized as follows. In section 2, we detail the calibration process we develop for laser-plane sensors located on CMMs. We show that the calibration process leads to a global transfer function that allows the transformation between the 2D data into 3D coordinates of points in the CMM workspace. The transfer function brings out parameters that must be identified. This point is exposed in section 3. The inverse problem is solved by the measurement of artefacts, the shapes of which are clearly defined and well known in the 3D workspace. However, some configuration parameters, such as the point position in the CCD space in the may affect measurements. The study of the most adapted artefact is thus conducted. Section 4 is dedicated to the development of the calibration algorithms for the chosen artefact, the facet sphere. In section 5, uncertainties associated with the method, in particular those linked to the variability of the parameters are estimated. This highlights the importance of the location in the 2D CCD space. The conclusion in section 6 indicates advantages and disadvantages of the proposed method.

2. Calibration process

The calibration process is the main element of an acquisition system for it allows the link between the 2D coordinates, $N(R, C)$, (R : row, C : column), in the CCD space to the 3D ones $M(X, Y, Z)$ to be expressed.

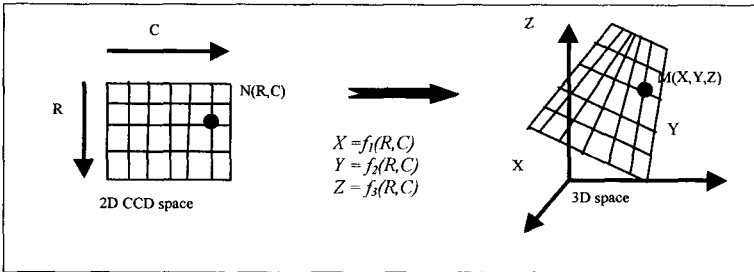


Figure 2. Calibration process: transformation of the 2D data into 3D coordinates.

Following the definition of Tsai⁶, “camera calibration in the context of 3D machine vision is the process that determines the internal geometric and optical

camera characteristics (intrinsic parameters) and/or the 3D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters)". Basically, the calibration process requires the definition of a calibration model, and procedures and algorithms to identify both extrinsic and intrinsic parameters⁷. The method we develop allows the identification of all parameters at the same time. Bases of the method that was first exposed by Dantan⁸ are developed next. They rely on the geometric modelling of the acquisition system.

2.1. Modelling of the acquisition system

Modelling requires the definition of coordinate systems attached to the different elements of the acquisition system⁸.

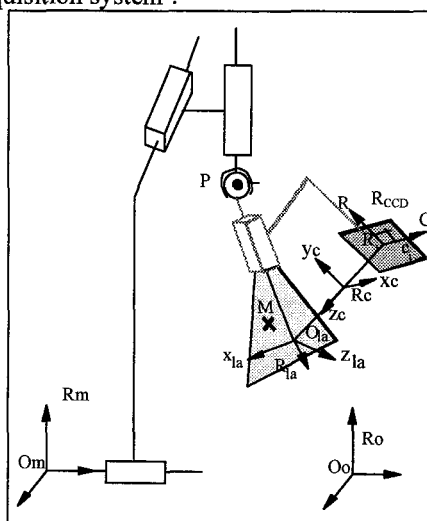


Figure 3. Relative positions and orientations of the different elements of the acquisition system.

CMM coordinate system: $R_m (O_m, X_m, Y_m, Z_m)$

R_m is associated with the CMM; its axes are the machine translation axes, and its origin is O_m , which is unknown.

Reference coordinate system: $R_o (O_o, X_m, Y_m, Z_m)$

This system is the absolute framework and is attached to a reference object located in the workspace of the CMM. The origin O_o is a characteristic point of the object.

CCD coordinate system: R_{CCD}

It is the coordinate system attached to the CCD camera, for which the axes R and C are parallel to the lines and columns of the CCD matrix. Coordinates of a point are referred to by (R_i, C_i) .

Camera coordinate system: $R_c (O_c, X_c, Y_c, Z_c)$

The origin O_c is the projection centre. The axis Z_c is merged into the optical axis. The axes X_c and Y_c are the projections of (R, C) of the CDD matrix into the plane normal to Z_c .

Laser-plane coordinate system: $R_{la} (O_{la}, X_{la}, Y_{la}, Z_{la})$

The origin O_{la} is the intersection of the optical axis of the lenses and the laser plane. The Z_{la} axis is normal to the laser plane and, X_{la} and Y_{la} are the projections of the directions (R, C) into the laser plane.

Let us consider a point M belonging to the object surface. The 3D coordinates (X, Y, Z) of M are expressed in the reference coordinate system R_o . M also belongs to the laser-plane space. Its coordinates in R_{la} are (X_{la}, Y_{la}, Z_{la}) . Therefore, we can write:

$$O_o M = O_o O_m + O_m P + P O_{la} + O_{la} M \quad (1)$$

$O_m P$ corresponds to the position of the sensor, which is identified by the CMM coordinates (reading of the measuring rules); $O_m P = X_s X_m + Y_s Y_m + Z_s Z_m$. $O_o O_m$ and $P O_{la}$ are translation vectors, which are a priori unknown. Let us consider R , the rotation matrix between the bases $B_m(X_m, Y_m, Z_m)$ and $B_{la}(X_{la}, Y_{la}, Z_{la})$; each vector U is transformed as follows: $U^{B_m} = R, U^{B_{la}}$ with

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

This leads to the following equation:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} T_{Omx} + T_{Olox} + X_s \\ T_{Omy} + T_{Oloy} + Y_s \\ T_{Omz} + T_{Olaz} + Z_s \end{bmatrix} + \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} X_{la} \\ Y_{la} \\ Z_{la} \end{bmatrix} \quad (2)$$

Eq. (2) brings out 6 parameters corresponding to the location of the sensor within the 3D space: 3 translations and 3 rotations. These parameters are generally called extrinsic parameters. However, most generally the rotation

movements are limited to 4 predefined positions of the sensor, considering the laser track parallel to one displacement axis (X-axis or Y-axis). Only the translation parameters have to be identified, which is simply carried out by the measurement of a sphere located within the workspace.

To express the global transfer function that links the 2D coordinates (R, C) to the 3D coordinates (X, Y, Z), we have to consider the transformation between the laser-plane system and the CCD coordinate system. We thus need to define the geometric model of the CCD camera.

2.2. Modelling the CCD Camera

The modelling consists of expressing the geometric behaviour of the CCD camera in the process of image formation through parameters. The most used model of the camera is the *pinhole model*. The camera objective is assumed to be an ideal lens, which is supposed to be plane (denoted the *image plane*), and the axes of the image plane are supposed to be perpendicular⁷. This model corresponds to a central projection, for which the centre of projection is the optical centre. Therefore, a point is projected onto the image plane by means of the optical centre. The distance between the image plane and the optical centre is the focal distance f .

Basically, the pinhole model leads to an analytical relation between the coordinates of the point expressed in the object space and those in the image plane:

$$(R, C) = F(X_c, Y_c, Z_c) \quad (3)$$

It can be noticed that an image point does not correspond to a single point, but to a single beam in the object space.

Parameters of the function F correspond to the intrinsic parameters associated with the CCD camera. This model can be integrated with effects linked to distortions taking into account various kinds of defects: geometrical quality of the lenses, misalignment of the lenses, This increases the number of parameters to be identified⁹.

2.3. Modelling in relation with the laser plane

The system that makes up the laser beam is composed of a light diode, a lens and a set of mirrors. The light diode is fixed on the optical axis. From a light source, the double lens will form a plane beam that focuses at a given distance.

The geometric model of the laser beam can be a plane, or a part of cone, or a polynomial surface.

As for the model of the CCD camera, this leads to a relationship between the (R, C) coordinates and the (X_{la}, Y_{la}, Z_{la}) coordinates related to the laser-plane coordinates through the central projection. Note that the model must take into account the angle between the laser-plane and the optical axis ⁶:

$$(R, C) = G(X_{la}, Y_{la}, Z_{la}) \quad (3)$$

For the 3D sensor used (KLS51, from Kréon Technologies), the laser beam is assumed to be plane. Therefore, as a 3D point also belongs to the laser plane (see Figure 2), its coordinate Z_{la} satisfies $Z_{la} = 0$ and Eq. (3) given by the constructor is ¹⁰:

$$\begin{cases} X_{la} = \frac{b_1 R + b_2 C + b_3}{d_1 R + d_2 C + 1} \\ Y_{la} = \frac{b_4 R + b_5 C + b_6}{d_1 R + d_2 C + 1} \\ Z_{la} = 0 \end{cases} \quad (4)$$

Eq. (4) brings out 8 additional parameters. These intrinsic parameters are linked to the sensor architecture and correspond to the geometrical modelling of both the CCD camera and the laser beam, which is assumed to be plane. Basically, these parameters are identified independently, most generally by the sensor constructor. Salvi¹¹ provides a survey of various camera-calibrating methods, highlighting the role of the chosen model and the identification method in the measurement uncertainty. Indeed, the accuracy of the camera calibration is one of the most important factors that condition measurement quality. However, the fact that geometrical parameters of the sensor are identified independently from the displacement system increases the measurement uncertainty.

2.4. Expression of the global transfer function of calibration

Let us now go back to the definition of a *global transfer function*. Our objective is to merge Eq. (2) and Eq. (4) so that we can directly express the coordinates of a 3D point in function of the position of the point in the 2D CCD space (R, C) . This leads to:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} + \begin{bmatrix} r_{11}X_{la} + r_{12}Y_{la} \\ r_{21}X_{la} + r_{22}Y_{la} \\ r_{31}X_{la} + r_{32}Y_{la} \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} + \begin{bmatrix} r_{11} \frac{b_1R + b_2C + b_3}{d_1R + d_2C + 1} + r_{12} \frac{b_4R + b_5C + b_6}{d_1R + d_2C + 1} \\ r_{21} \frac{b_1R + b_2C + b_3}{d_1R + d_2C + 1} + r_{22} \frac{b_4R + b_5C + b_6}{d_1R + d_2C + 1} \\ r_{31} \frac{b_1R + b_2C + b_3}{d_1R + d_2C + 1} + r_{32} \frac{b_4R + b_5C + b_6}{d_1R + d_2C + 1} \end{bmatrix}$$

Regrouping in the previous equation terms in R and C yields

$$\begin{cases} X = X_s + \frac{a_1R + a_2C + a_3}{a_{10}R + a_{11}C + 1} \\ Y = Y_s + \frac{a_4R + a_5C + a_6}{a_{10}R + a_{11}C + 1} \\ Z = Z_s + \frac{a_7R + a_8C + a_9}{a_{10}R + a_{11}C + 1} \end{cases} \quad (5)$$

The global transfer function of calibration (Eq. (5)) expresses the coordinates (X, Y, Z) of a 3D point in the CMM-system as a function of the location (R, C) in the CCD space of the sensor. This expression brings out 11 global parameters, which are representative of both the intrinsic, and the extrinsic parameters. The calibration of the geometrical parameters of the sensor and the calibration of the sensor location within the workspace are performed through the same identification process. Uncertainties are thus limited. All sensor orientations can be considered.

The identification process relies on measurements of artefacts of specific shapes. The choice of the artefact shape is important for the efficiency and the robustness of the calibration method.

3. Identification of calibration parameters

The identification process entails the study of the geometry of the artefact, which has to be simple and reachable whatever the orientation of the sensor. This artefact becomes the reference element. The shape is conditioned by the following constraints:

- geometric model,

- accessibility of the sensor.

The most usual artefacts used in dimensional metrology are spheres and prismatic elements, for their geometry is simple. Furthermore, fitting algorithms of an ideal geometrical element to the data points are well controlled¹².

However, in this case the problem is quite different. Our data are the coordinates (R_i, C_i) in the 2D space of an object which position is known in the 3D space. Therefore, considering eq.5, for a given sensor orientation, the identification of the 11 parameters relies on the measurement of 11 points M_k well known in the 3D space. In practice, it is difficult to isolate a point in the CCD space. As a measure corresponds to the intersection between the laser-plane and the surface object, the result in the 2D space is a *digitising line* (see Figure 4).

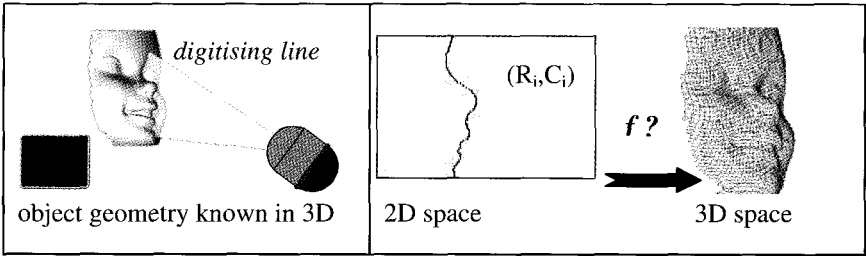


Figure 4. Measure in the 2D space: *the digitising line*

First, let us consider that the identification is performed through the measurement of a sphere. Let us suppose that another measuring means, a contact probe, has precisely measured the sphere. Let C be the centre of the sphere and R be its radius.

Each point M_k , in the 3D space, belonging to the sphere satisfies:

$$(X_k - C_x)^2 + (Y_k - C_y)^2 + (Z_k - C_z)^2 = R^2$$

Each point M_k that has been measured also satisfies Eq. (5). That means that there exists at least one point $N(R_j, C_j)$ corresponding to M_k in the 2D

$$\begin{aligned} \text{space:} \quad & (X_s + \frac{a_1 R_j + a_2 C_j + a_3}{a_{10} R_j + a_{11} C_j + 1} - C_x)^2 + (Y_s + \frac{a_4 R_j + a_5 C_j + a_6}{a_{10} R_j + a_{11} C_j + 1} - C_y)^2 \\ & + (Z_s + \frac{a_7 R_j + a_8 C_j + a_9}{a_{10} R_j + a_{11} C_j + 1} - C_z)^2 = R^2 \end{aligned} \quad (6)$$

In Eq. (6), the linearity between parameters is lost. This complicates the identification problem. However, measuring a large number of M_k points (more

than 11) could solve the problem. Furthermore, we have to define a measurement protocol in order to be sure that all the M_k points are sufficiently representative of the 3D geometry of the sphere. Note that, as the orientation of the sensor must be preserved during the whole calibration stage, only transformations that leave the laser plane invariant are thus possible. That means 2 translations in the plane (X_{la} , Y_{la}) and 1 rotation around Z_{la} .

Considering the previous remarks, the plane seems an interesting reference element, as it satisfies a linear equation. However, a single plane does not allow the calibration of all sensor orientations. This leads us to define a specific artefact that consists of various planes: *the facet sphere* (Figure 5).

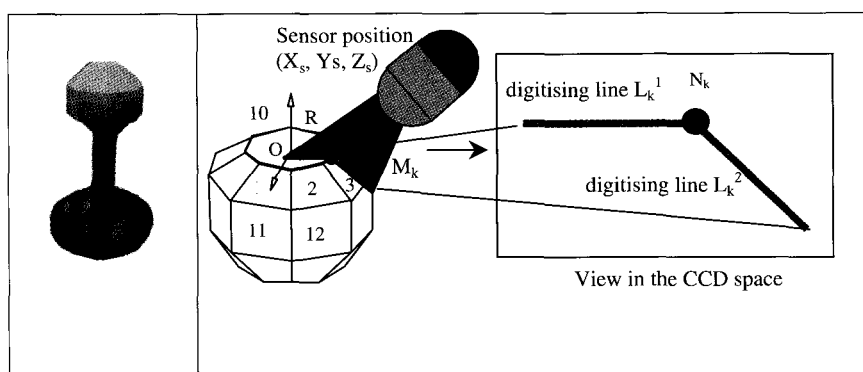


Figure 5. The facet sphere: measurement of a 3D point.

As we have to measure at least 11 points in 3D to identify the 11 parameters, we choose to define interest points. These points are located on an edge of the artefact, and correspond to the intersection between two adjacent planes. In the 2D space these points result from the intersection between 2 digitising lines. As they are seldom measured, they are defined by calculation. Algorithms we developed relative to the calibration process based on the facet sphere are presented in the next section.

4. Calibration process using a specific artefact: the facet sphere

The equation of calibration (Eq. (5)) brings out 11 parameters, a_i ($i = 1, \dots, 11$) that are identified by the extraction of interest points measured on 3 adjacent facets. The facet sphere is geometrically well known and located within the workspace of the CMM (measured using a mechanical probe). All the plane faces have a form deviation less than $1 \mu\text{m}$. To be measured via the KLS51

sensor, all faces must be coated with a white powder, which unfortunately increases the digitising noise. The faces are numbered from 1 to 25, and the equation of the corresponding plane, in the CMM workspace, is $\alpha_j x + \beta_j y + \chi_j z = \gamma_j$, with $j \in [1, 25]$.

As previously indicated, the calibration process consists of the measurement of interest points, M_k . These points are located on the artefact; $M_k(X_k, Y_k, Z_k)$ in the CMM space. They are observed in the 2D space; that means that there exists a point $N_k(R_k, C_k)$ corresponding to M_k . However, as the point belongs to the intersection between two adjacent planes, N_k is the intersection between two digitising lines, L_k^1 and L_k^2 , traces of the planes in the 2D space (Figure 5).

Therefore, M_k satisfies, in the CMM space, the equation of each plane to which it belongs:

$$\alpha_j X_k + \beta_j Y_k + \chi_j Z_k = \gamma_j \quad (6.a)$$

$$\alpha_p X_k + \beta_p Y_k + \chi_p Z_k = \gamma_p \quad (6.b)$$

M_k also satisfies the global transfer function and by substitution of (X_k, Y_k, Z_k) given by Eq. (5) into Eq. (6.a) and (6.b), the following two equations that are linear relative to the unknown parameters a_i are obtained:

$$\begin{aligned} -K_j &= a_{11}K_jC_k + a_{10}K_jR_k + a_9\chi_j + a_8\chi_jC_k + a_7\chi_jR_kC_k + a_6\beta_j + a_5\chi_jC_k \\ &+ a_4\beta_jR_k + a_3\alpha_j + a_2\alpha_jC_k + a_1\alpha_jR_k \end{aligned}$$

$$\begin{aligned} -K_p &= a_{11}K_pC_k + a_{10}K_pR_k + a_9\chi_j + a_8\chi_jC_k + a_7\chi_jR_kC_k + a_6\beta_j + a_5\chi_jC_k \\ &+ a_4\beta_jR_k + a_3\alpha_j + a_2\alpha_jC_k + a_1\alpha_jR_k \end{aligned}$$

where $K_r = \alpha_r X_s + \beta_r Y_s + \chi_r Z_s + \gamma_r$ ($r = j$ or p). Remember that the coordinates (X_s, Y_s, Z_s) corresponding to the linear position of the sensor within the CMM space are given by the measuring rules.

Each acquisition of a point of interest leads to 2 linear equations, and the parameters are the solution of a linear system of $2N$ equations in 11 unknowns, where N is the number of points of interest, allowing the identification. For robustness reasons, the resolution is performed using a superabundant system of 36 equations (that corresponds to 18 acquisitions of various points of interest). The choice of these 18 acquisitions results from a specific measurement protocol for the calibration. Indeed, as said previously, we must ensure that all points are distinct within the CCD space.

The protocol is thus the following: let us choose 3 parallel sections, for instance (face 10, face 3), (face 10, face 2), and (face 10, face 1); for each section, we move the sensor following 2 heights and 3 sweepings within the

laser plane while preserving the sensor orientation. For each position, coordinates of the sensor are acquired and the point of interest is calculated in the CCD space by the intersection between the two digitising lines (Figure 6). The linear system is then solved (using the Least-Squares criterion) giving the value of the 11 parameters a_i constituting the global transfer function f .

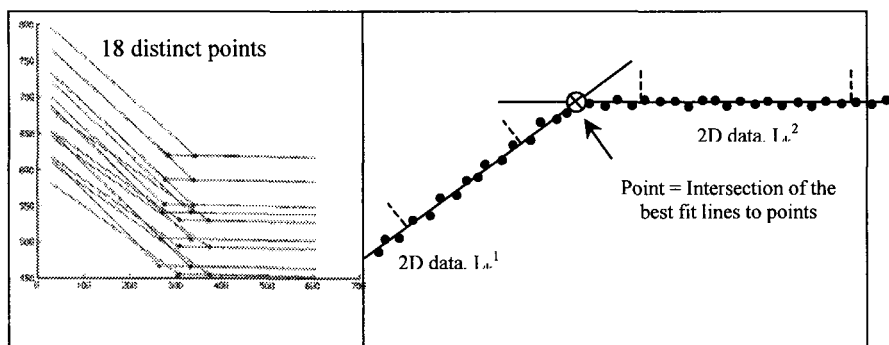


Figure 6. Acquisition of the points of interest in the 2D space.

However, the points of interest result from the calculation of the intersection between 2 digitising lines. As the 2D data are largely noisy, we associate the best *least squares line* fitting the data to carry out the intersection calculation. This method of calculation is itself a source of errors. Indeed, the association is performed on a portion of the digitised data for each line. This portion allows border effects to be limited, in particular those near the measured edge. Due to the digitising noise, another measurement gives a different set of data points. The equation of the line fitting the points is thus different from the previous one. Therefore, there is an uncertainty of the position of the interest point, which is of great influence on the identification of the global transfer function. This point is discussed in the next section.

5. Uncertainty linked to the calibration method

The digitising noise may come from various sources^{5,13}: laser influence, characteristics of the employed optics, behaviour of the CCD camera, aspect of the object surface, ... This point is not discussed here, but, whatever its source, we suppose that the digitising noise exists, and is a factor influencing the accuracy of the calibration method. We can add to the influencing factors the anisotropic behaviour of the CCD matrix: the acquisition is strongly related to the observation location in the 2D space of the CCD camera.

As said previously, a point of interest results from the intersection of two lines, each one corresponding to the intersection between the laser plane and a specific face of the facet sphere. The accuracy of the identification of both lines depends on the digitising noise, and impacts the global transfer function. Indeed, a first calibration procedure relying on the 18 calculations of N_k leads to the set of parameters $(a_i)_1$. However, due to the digitising noise, a new calibration procedure of the system may provide another set of parameters $(a_i)_2$, and so on. Each set of parameters introduced in Eq. (5) leads to different values of (X, Y, Z) , for the same location (R, C) in the 2D frame, and involves a dispersion on the location of the 3D point.

$$\begin{cases} X + \Delta X = X_s + f_1(R, C) + \Delta f_1 \\ Y + \Delta Y = Y_s + f_2(R, C) + \Delta f_2 \\ Z + \Delta Z = Z_s + f_3(R, C) + \Delta f_3 \end{cases} \quad (7)$$

To evaluate the dispersions due to the extraction of the points of interest and the impact $(\Delta X, \Delta Y, \Delta Z)$ of such dispersions on the obtained 3D points, we develop a method relying on numerical simulations.

A first calibration procedure is performed giving 18 acquisitions, which constitute the reference. From this first step, digitising noise must be evaluated. An indicator of noise δ is calculated. Considering that a line is fitted to the noisy data according to the least-squares criterion, the standard deviation, σ , of the deviations of the data relative to the line is calculated (Figure 7b). For each of the k acquisitions L_k^j (k varying from 1 to 18), we associate the noise δ_k^j , considering that $\delta_k^j = \sigma_k^j$ (with $j = 1$ or 2 for each one of the lines). Note that only a subset of n_k^j points is preserved in order to be free from inaccuracies due to edge effects (ends of the laser line or sharp corner) (Figure 7a).

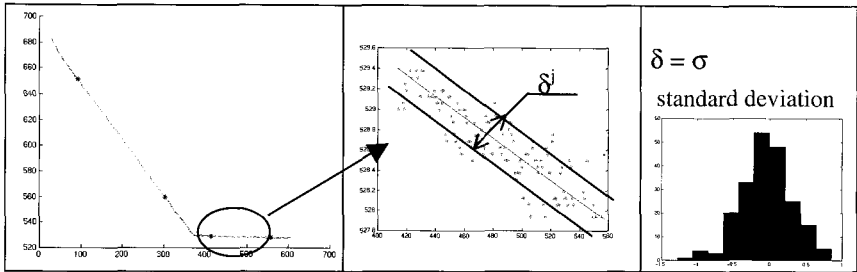


Figure 7. Evaluation of the digitising noise for each acquisition.

Then we carry out the numerical simulations that correspond to virtual calibration procedures¹³. The n_k^j points are randomly scattered about the initial

line, according to the noise δ_k^j . To each new point distribution a straight line is fitted using the least-square criterion. The identification of each of the N_k points of interest is carried out by the calculation of the intersection of the least-squares simulated lines. For each acquisition k , we perform N_b simulations; this gives N_b points N_k ($k=1, \dots, 18$). By reporting the results in Eq. (5), we obtain various sets of the 11 parameters $(a_i)_p$, for $p = 1, \dots, N_b$.

The impact of the digitising noise on the location of a 3D digitised point is calculated as follows. A grid of points is generated in the 2D CCD frame (Figure 8a). For each set of parameters $(a_i)_p$, we perform the transformation of the 2D grid points into 3D points, using the global transfer function (5). This operation is repeated for $p = 1$ to N_b . As a result, to each point of the 2D grid correspond N_b points, defining a zone. This zone corresponds to the dispersion of the acquisition of a 3D point, a dispersion that takes into account the effect of the digitising noise on the calibration procedure (Figure 8c).

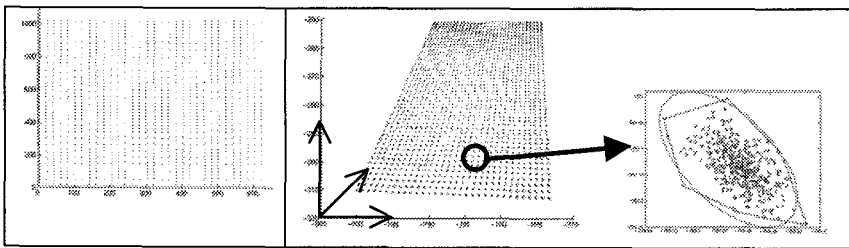


Figure 8. Simulation of the 3D space according to Eq. (7) – Dispersion of the location of a 3D point.

The value of the dispersion can be quantified by the value of the equivalent disc area of the ellipse that fits the point set. The obtained 3D-uncertainty zone is a function of the acquisition location in the 2D space, which confirms the anisotropic behaviour of the CCD matrix. Results show an extension of the position uncertainty from 0.05 mm to 0.3 mm in peripheral zones of the laser plane seen by the CCD camera.

6. Conclusion

In the paper we proposed a method for the calibration of laser-plane sensor on CMMs through the definition of a global transfer function. The method allows the transformation of 2D data into 3D coordinates in the CMM system for all orientations of the sensor supported by the CMM. The transfer function brings out 11 parameters that are identified at the same time by measurement of a specific artefact, the facet sphere. Little work has yet been done in this direction.

However, as for any type of calibration method, the process is a source of dispersion. Indeed, due to the digitizing noise, there is variability on the parameters of the transfer function. The simulation we propose defines virtual calibrations, and leads to an estimation of the parameters variability. It is thus possible to evaluate the impact on the uncertainty of measured points by optical measuring system. This highlights that uncertainty with such measuring systems is important, and cannot allow precise measurements for dimensional metrology.

References

1. P. Bourdet, C. Lartigue and F. Leveaux, Effects of data point distribution and mathematical model on finding the best-fit sphere to data, *Precision Engineering*, vol. 15, n°3, 150-157 (1993)
2. S. Sartori, Geometric Error Measurement and Compensation of machines, *Annals of the CIRP*, vol. 44, n°2, 599-609 (1995)
3. P. Bourdet, Contribution à la mesure dimensionnelle: modèle d'identification des surfaces, métrologie fonctionnelle des pièces mécaniques, correction géométrique des machines à mesurer, *Thèse d'Etat, Université Nancy I, France* (1987)
4. J. Salvi, X. Armang and J. Batlle, A comparative review of camera calibrating methods with accuracy evaluation, *Pattern recognition*, vol. 35, 1617-1635 (2002)
5. A. Contri, P. Bourdet, C. Lartigue, Quality of 3d digitised points obtained with non-contact optical sensor, *Annals of the CIRP*, vol.5, n°1, 443-446 (2002)
6. R. Y. Tsai, A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses, *IEEE Journal of robotics and automation*, Vol. RA-3, n°4 (1987)
7. J. Zhou, Contribution aux méthodes d'étalonnage des capteurs d'images, *Thèse de doctorat, Ecole Centrale Paris / ISMCM, France* (1992)
8. J.Y. Dantan, R. Rey, P. Bourdet, D. Feautrier, Calibrating the geometric position of a plane laser-beam vision sensor in a measuring system for tridimensional shapes, *Proceedings of International Mechatronics Congress, Besancon (France)* (1996)
9. J. Weng,, P. Cohen, M. Herniou, Camera Calibration with distortions models and accuracy evaluation, *IEEE Transactions on Pattern and Machine Intelligence*, vol. 14, n°10 (1992)
10. M. Brunet, E. Cosnard, B. Deltour, Procédé d'étalonnage d'un système d'acquisition tridimensionnelle de forme et du système pour la mise en oeuvre du dit procédé, Brevet 04 52 422 B1, KREON Ind., (1993)

11. J. Salvi, X. Armangué, J. Battle, A comparative review of camera calibrating methods with accuracy evaluation, *Pattern recognition*, vol. 35, 1617-1635 (2002)
12. A. Forbes, Robust circle and sphere fitting by least-squares, *NPL Report DITC 153/89* (1989)
13. A. Contri, Qualité géométrique de la mesure de surfaces complexes par moyens optiques, *PhD thesis*, ENS de Cachan(France), (2002)

SOME DIFFERENCES BETWEEN THE APPLIED STATISTICAL APPROACH FOR MEASUREMENT UNCERTAINTY THEORY AND THE TRADITIONAL APPROACH IN METROLOGY AND TESTING

C. PERRUCHET

Chairman ISO/TC 69

UTAC

BP212

F-91311 Montlhéry cedex, France

E-mail: christophe.perruchet@utac.com

Statistical methodology coming from the mathematical branch of probability theory, and powerfully helped today by the alliance of computing science and other fields of mathematics gives ways of "making the data talk". The place of statistical methods in metrology and testing is presented. Emphasis is put on some difficulties encountered when wishing to apply statistical techniques in metrology: importance of the traditional mathematical assumptions of many statistical techniques, difference between a confidence interval and other intervals (coverage interval,...), difficulties caused by the differences in terminology and concepts used by metrologists and statisticians. At last information is given about standardization organizations involved in metrology and statistics.

1. The place of Statistics

A key role is played today by the production, collection, analysis, presentation and interpretation of data. But it is useless to collect data if it is not to be analyzed and interpreted with a view to enlighten human actions or to progress the knowledge of phenomena. In summary, the aim is to transform raw data, once collected, into usable, understandable and communicable information.

Statistical methodology coming from the mathematical branch of probability theory, and powerfully helped today by the alliance of computing science and other fields of mathematics (linear algebra, graph theory, algorithmic...), gives ways of "making the data talk".

Different perspectives may be used to see what today is the place of statistics.

The first perspective "from facts to decision" is to place statistics in a classic process starting with facts and ending with actions and decisions.

Such a process is described by the sequence:

Events/facts -> Data -> Information -> Knowledge -> Goals -> Action plans ->
Actions/decisions -> Events/facts -> ...

Statistical techniques such as design of experiments are used in the "events/facts subprocess". Data collection and data organization techniques are used in the

"data" subprocess. Data analysis techniques are used in the "information" subprocess. Statistical decision theory techniques are used in the "action plans" and "actions/decisions" subprocesses.

Another interesting perspective is to place statistics in the "information management cycle" inspired by Deming's PDCA.

The "plan" stage is about data planning. It is the domain of planification where statistics are used for the data design, including available data and data to be produced, and the data processing design.

The "do" stage is about data production and management. It is the domain of data management where statistics are used for data collection, data checking, data organization and management.

The "check" stage is about the production of information and knowledge. It is the domain of knowledge where statistics are used for data access, off line or on line data analysis and reporting.

The "action" stage is about decision making. It is the domain of action and decision where statistics are used for decision rules and indicators.

A third perspective is to use the life cycle of products (see [1]). The starting point and ending point of this cycle are respectively customer expectations and customer satisfaction. The megaphases of this cycle are conception, development, delivery, maturity and death.

The concept of variation, and the need to interpret data, and filter attributable effects from "noise" is inherent in all phases of the product life cycle. Statistical methods are used in the collection and interpretation of data on customer needs. The use of the methods of experimental design is fundamental to the product and process development and design phase. Statistical methods for process analysis process control, and process improvement are needed to monitor and improve the process performance. Reliability methods and statistical models for condition monitoring and analysis of failure data are of importance for monitoring product and process during the operational lifetime.

From this short presentation it appears that statistics is not only a set of techniques but is a methodology for collecting data, analysing data, transforming data into information and designing decision rules, which takes place in the product life cycle and in the information management cycle. More information can be found in [2].

2. Issues in Metrology

Traditionally, in metrology related activities, statistical methods have been utilised and referred to in the description of measurement and test methods for properties of product and in sampling inspection. The need for application of statistical methods is, however, not limited to these activities.

Statistical methods are needed for assessing measurement uncertainty, for calibration, monitoring and improving the measurement processes at the producer's site, as well as being needed by the various agencies involved in testing, verifying conformance, and validating the producer's quality and environmental management systems.

However, to improve the role statistics may play for metrology, it is important to identify issues on which cooperation between experts of these domains has to be reinforced.

2.1. Terminology and concepts

Using consistent concepts is of major importance. Without pertinent concepts the risk is to all agree if one says that the right angle boils at 90° .

Two reference documents have to be used in the domains of metrology and statistics. These documents are ISO standards: the VIM [3] for metrology and ISO 3534 [4] for statistics. Of course these two systems of concepts do not cover the same domain, their interaction is on concepts related to applied statistics in metrology and testing activities.

Unfortunately, in this field of interaction differences exist. As a consequence, practical difficulties for laboratories come from these discrepancies, either on concepts themselves, either on the wording of the concepts' definition.

Two typical examples can be given.

In the VIM the concept of error relates to the deviation between a data (measurement result, indication...) and a "true" value. In ISO 3534, the concept of (statistical) error, also called residual, relates to the deviation between a data and the (statistical) model.

In the VIM the concept of repeatability relates to a situation where all conditions of measurement are the same. In ISO 3534, the concept of repeatability relates to the sources of variation which are not explicitly taken into account in the model and which therefore are included in the variance of the residual (the error term of the statistical model).

2.2. Usual assumptions and robustness

Another issue is the importance of the usual assumptions made when using a statistical model. In general, the two main assumptions are normality and independence.

A statistical model has to be seen under two perspectives. The first one is the fit to the input data. This is a geometric perspective (how does the model fit the data?) in which no assumption on the data distribution (e.g. normality) is necessary, as long as no inference is made.

The second perspective is related to the use of the model to infer conclusions to another data or situations. This is why assumptions have to be made on the data distribution. It is well known that the mathematical properties of the normal

(gaussian) distribution allowed for developing most of the models which are practically used.

However, many "natural" phenomena are not "normal" (i.e. distributed according to the gaussian law).

Working with low levels of pollutants implies to take into account asymmetry and truncation. Measuring the largest diameters of a set of rods implies to work with extreme values distribution. Analysing failure occurrences imply to use statistics of rare events. Reliability tests imply to work with censored data.

Today, efficient solutions for these "non normal" situations are given by numerical simulation techniques. These techniques, such as Monte-Carlo simulation or bootstrap, allow taking into account the "true" distribution of observed data. In this respect, however, the quality of data deserves careful thinking (Garbage In Garbage Out).

The second assumption used by statistical models is independence. Actually, non correlation is often a sufficient assumption as long as second order moments (variances) are considered.

Two kinds of independence have to be considered. The first kind is independence between the random variables, present on the right side of the equation, which represent the explanatory variables (the so-called "independent variables") and the residual. The second kind is the independence between the random variables representing repeated measurement.

The question of testing or checking for this assumption is still very difficult to address. One of the reasons why is that different concepts are mixed: stochastic independence for the statistical model, statistical correlation generally used as the tool for evaluating non correlation and "substantive correlation" as evaluated from prior knowledge, not directly related to the data used, for evaluating the statistical model. This question may be considered as remaining open.

2.3. *p* values

The users' meanings of the null hypothesis of significance tests cover a very wide range. Users have also major difficulties for interpreting *p*-values. Apart from everyone's experience, these have been reported in different publications (see, e.g. [5]).

Actually, a major confusion exists between statistical significance, expressed in terms of *p*-values and "substantive significance" related to the expert knowledge of the studied phenomena. This should be improved by using more statistical thinking and less statistical rituals.

A consequence of these confusions and of the difficulties in verifying the model assumptions is the poor confidence that can be put in the traditional values of the coverage factor *k* and the coverage probability *p* in measurement uncertainty.

These traditional couples of values for (*k*, *p*) of (2, 0.95) and (3, 0.99) have to be considered only as conventional values (see, e.g. [6]). Unless when

assumptions can be met or when a careful numerical simulation is performed, the value of p , as a function of k , can be estimated with a good level of “confidence”. Moreover, it would be preferable to derive a pertinent value of p from a risk analysis of the intended use of the uncertainty of measurement, and estimate a value of k from the value of p .

2.4. Intervals

Confusions and difficulties also occur with the different kinds of intervals that may be used. Apart from difficulties in understanding the differences between the concepts this is also caused because their names are formed from a limited number of words such as: interval, confidence and coverage.

A coverage interval is defined (GUM §2.3.5) as an “interval ... that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand...Note 1: the fraction may be viewed as ... the level of confidence of the interval.”

A confidence interval is defined (ISO/CD 3534-1.3 §3.28) as an “interval estimator ... for the parameter...with the statistics ...as interval limits...Note 1: The confidence reflects the true proportion of cases that the confidence interval would contain the true parameter value in a long series of repeated random samples (3.6) under identical conditions. A confidence interval does not reflect the probability (4.6) that the observed interval contains the true value of the parameter (it either does or does not contain it).”

A statistical coverage interval is defined (ISO/CD 3534-1.3 §3.26) as an “interval determined from a random sample (3.6) in such a way that one may have a specified level of confidence that the interval covers at least a specified proportion of the sampled population (3.1) ... Note: The confidence in this context is the long-run proportion of intervals constructed in this manner that will include at least the specified proportion of the sampled population.”

It is important to notice that these concepts are different and that the whole system is consistent.

This means that a coverage interval is not a confidence interval. A coverage interval does not estimate a parameter and the limiting values are not statistics (when type B evaluation is used).

A coverage interval is not a statistical coverage interval either.

It also has to be noticed that in the definitions of confidence interval and statistical coverage interval, the level of confidence is properly defined in a “frequentist way”.

This is not the case in the definition of the coverage interval. It is necessary to go to other parts of the GUM to see that the level of confidence is rather a degree of belief.

Therefore, what can be the users’ meaning of “confidence”? Efforts should certainly be undertaken for clarifying these concepts in the users mind.

2.5. Use of measurement uncertainty

Laboratories devoted many resources in past years for estimating their uncertainties of measurement. These efforts have to be continued as measurement equipment and methods are evolving. However, today's issues are also about the use of measurement uncertainty.

Three main domains can be identified where methods describing how to use measurement uncertainty should be improved or developed.

In conformity assessment (of a product) a decision is taken, from a sample (item), about the conformity of a population (lot). In related international standards two sources of variation are today taken into account in: sampling and variability of production. Measurement uncertainty, as representing the uncertainty of the measurement process, is generally not explicitly identified as a source of variation. When it is identified, it is assumed as being negligible compared to the other sources of variation.

However, situations exist where this is not the case, e.g. for chemical analysis, it is therefore necessary to provide for methods taking into account all the relevant sources of variation.

A second issue is the evaluation of the capability of equipment and processes. The capability characterises the degree to which equipment or processes respect specifications. This concept can also be applied to test and measurement methods. Multiple definitions of this concept, coming from different fields of activity, are today available. Measurement uncertainty should always be explicitly taken into account.

A crucial question is the control of measurement uncertainty over time. In a typical evaluation of measurement uncertainty, time does not appear as a source of variation, as it is implicitly assumed that measurement uncertainty is not time dependent. When capability evaluation is considered it is necessary to consider the stability of the processes over time; therefore the dependency of measurement uncertainty over time has to be explicitly considered.

The last issue is the evaluation of the proficiency of laboratories. Collaborative or interlaboratory studies have been proved to be a powerful tool for this purpose. A recent ISO standard, ISO/DIS 13528 [7], introduces the use of measurement uncertainty in proficiency evaluation. However, measurement uncertainty is not the only parameter to take into account. In this respect it is necessary that all the information used related to the quality of the information processed (laboratories' results, reference values,...) be established on a consistent basis.

2.6. Metrology and Testing

It is important to understand that metrology and testing are domains which are not identical even if they largely overlap.

The purpose of a measurement is to evaluate a measurand; the purpose of a test is to evaluate a characteristic.

A measurand (i.e. a particular quantity subject to measurement) is defined in consistency with the SI system of units and is independent of the measurement method, in the sense that a given quantity can be determined by using different methods of measurement.

On the contrary, a characteristic do not necessarily refer to the SI system of units (e.g. hardness, viscosity, sweetness,... and other words ending in "-ess") and is defined in the context of a given test method. The viscosities determined by different test methods are not the same viscosities (unless the "correlation" between test methods has been established).

As a consequence the sub-concepts which are related to the measurand cannot be related to the characteristics but have to be related to the test methods. An example is given in the difference in the definitions of reproducibility in VIM and in ISO 3534.

Moreover a characteristic do not necessarily takes its values in a continuous set (interval) of numerical values (real or complex, in dimension-1 or dimension-n). These are called ordinal characteristics. Examples are mark scales or ranks.

There are also characteristics for which the possible values are not numerical and do not have an order structure. These are called nominal characteristics. Examples are pass/fail test results for which there are two possible values for the test result (pass or fail) or more than two values (if the cause of the fail is recorded)

As a consequence pertinent statistical and data analysis techniques have to be used for these characteristics which do not pertain to traditional metrology.

3. Resources: standardization committees

The fundamental work of standardization organizations has to be pointed out. In a transversal domain such as metrology, efficient communication between experts of different fields (physics, chemistry, statistics, quality,...) is a necessary condition for progressing. International standards being documents on which a consensus has been reached it is therefore essential to improve the coordination between standardization organizations.

The situation is certainly better today than it was in the past, as a first coordination is ensured by the participation of some experts to several of these committees. This is the case e.g. for ISO/TC69 (the ISO technical committee developing standards on the application of statistical methods) and JCGM (under which are undertaken the revision of VIM and the drafting of supplemental guides to the GUM). This participation improves the consistency between documents that are under development.

However this operational coordination is not sufficient. It is now necessary to establish coordination at the strategic level in order to improve the coordination

of the business plans of these organizations. This should of course also include the activities of regional and national organizations.

More information on the work of ISO/TC 69 and JCGM can be found on the following web sites: www.iso.ch and www.bipm.org.

References

1. M. Boulanger, M. Johnson, C. Perruchet and P. Thyregod: Evolution of international statistical standards via life cycle of products and services, *Int. Stat.. Rev.* **Vol. 67-2**, 151-171 (1999).
2. C. Perruchet: Introduction to data mining, *Stat'Expo tutorial (slides)* (1997).
3. International vocabulary of basic and general terms in metrology (VIM), ISO (1993)
4. ISO 3534 Statistics: Vocabulary and symbols, ISO (1993)
5. B. Lecoutre, M-P Lecoutre, J. Poitevineau: Uses, abuses and misuses of significance tests in the scientific community: won't the Bayesian choice be unavoidable, *Int. Stat.. Rev.* **Vol. 69-3**, 399-417 (2001).
6. C. Perruchet, M. Priel: Estimer l'incertitude, Mesures, Essais, *Afnor* (2000)
7. ISO/FDIS 13528: Statistical methods for use in proficiency testing by interlaboratory comparison, ISO (2002)

METROLOGY SOFTWARE FOR THE EXPRESSION OF MEASUREMENT RESULTS BY DIRECT CALCULATION OF PROBABILITY DISTRIBUTIONS *

GIOVANNI B. ROSSI, FRANCESCO CRENNNA, MICHELE CODDA

University of Genova –DIMEC

*PMAR-Lab Laboratory of Design and Measurement for Automation and Robotics
Via All'Opera Pia 15 A, 16145 Genova*

In this paper the possibility of expressing the final result of a (any) measurement by a probability distribution over the set, discrete or continuous, of the possible values of the measurand is considered.

After a brief review of the motivation, state of the art and perspective of this option, the related software tools are discussed, with reference to the package UNCERT developed at the authors' laboratory.

The software implements an approach based on the direct calculation of probability distributions, for a wide class of measurement models. It is arranged in a hierarchical and modular structure, which greatly enhances validation.

Involved variables are considered as inherently discrete and related quantisation and truncation effects are carefully studied and kept under control.

1 Introduction

The possibility of expressing the final result of a (any) measurement by a probability distribution over the set, discrete or continuous, of the possible value of the measurand is a very attractive option for metrology. The reasons for that are manifold, concerning both the measurement evaluation (or restitution) process and the information communication, and may perhaps be summarised as follows:

1. the expression of the final result of measurement by a probability distribution is a highly satisfactory format from a theoretical standpoint, at least as soon as probability is accepted, as usually is, as an appropriate logics for dealing with uncertain events/statements [1];
2. it allows calculation of any statistics needed for application (expected value, standard deviation, intervals...[2-4]) including those related to weaker, for example ordinal, measurement scales, which are of increasing theoretical and practical interest [5];

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

3. the propagation of probability distributions may be a necessary approach to deal with those aspect of uncertainty evaluation “for which the GUM¹ is insufficiently explicit and for which it is not clear whether its conventional (mainstream) interpretation is adequate” [6], so that this topic is presently under consideration by the Working Group 1 (WG1), Measurement Uncertainty, of the Joint Committee for Guides in Metrology (JCGM) [7];
4. expressing the results of measurement in terms of probability is essential when risk and cost assessment are involved, and, in general, when measurement is intended to support critical decisions: anyone sees how these topics are of high socio-economic impact and are strategic for the measurement community [8-10].

Let us now briefly consider the problems associated with this option. They are of both theoretical and practical nature and include:

- the “meaning” of such a description (this is an essentially theoretical and even epistemological problem, but has also practical consequences);
- the “model” that may be assumed for supporting such a description;
- how to assign required probability distributions (those involved in the model);
- the supporting metrology software.

We are here mainly concerned with the last one, but we are convinced that any progress in this aspect will encourage pursuing this approach and will contribute to a progress in the theoretical debate. Software in support of a probabilistic expression may follow two main approaches, namely using stochastic simulation techniques [6] or performing direct calculation of involved probability distributions [11]. Although the first approach is at present more studied, we consider here the second one, since we are convinced of its validity. This includes the possibility of using it for checking the results obtained through simulation, through a back-to-back comparison strategy.

So the article is organized as follows: the main theoretical background is first concisely presented, including an outline of the assumed reference probabilistic model and the derivation of two specific models of wide interest. Then the development and implementation of the package UNCERT is addressed, starting from the hierarchy of models considered, going through some essential implementation details and considering quality-assurance aspects

¹ GUM is an acronym for the “Guide to the expression of uncertainty in measurement” [12]; for the “mainstream GUM” method, see section 2.2 of this article.

also. Finally, a number of illustrative test cases are discussed, showing the potentials of the package, and conclusions are drawn.

2 The probabilistic model

2.1 The probabilistic reference model

Modelling of measurement processes is a key topic of measurement theory [14], with several operational implications also [15-17]. Although we surely have no room here for even outlining it, we simply recall some general features of the model presented in Ref. [1], which we will use as our reference model in the following.

Let us then consider a measurable quantity x , with values defined in a set X , and the class of measurement tasks consisting in the measurement of x , under the assumption it has a constant value for an interval of time T , which contains the interval of time T_0 required for entirely performing the measurement, including eventual repeated observations.

We consider the overall measurement process as resulting from the concatenation of two sub-processes, named *observation* and *restitution*. This distinction is very important and completely general, since the former accounts for transformations in the measuring system giving rise to the observable output, the latter includes data processing which yields the final measurement result. The output of the observation process is in general a vector $y = [y_1, \dots, y_N] \in Y$ of observations. This formalism may represent measurement processes based on repeated observations from the same measuring system as well as indirect measurements, “in which the value of the measurand is obtained by measurement of other quantities functionally related to the measurand” [12].

We also introduce the vector of *influence parameters* θ , with values $\theta = [\theta_1, \dots, \theta_K]$ in a K -fold space Θ . Influence parameters are of two kinds:

1. *influence quantities*, i.e. “quantities which are not the subject of the measurement, but which influence the value of the measurand or the indication of the measuring instrument” [12];
2. parameters concerning the involved probability distributions, typically *dispersion parameters*.

Actually parameters of the second class may be thought as regarding a higher level of modelling, as a kind of hyper-parameters, but since they may be mathematically treated in the same way as those in the first class, we prefer to avoid such a complication.

We also assume a probability measure to be defined over Θ , so that the probability distribution $p(\theta)$ is defined. Then the *observation process* may be characterised by a parametrical probabilistic mapping from the set of values of the measurand X to the set of the observations Y , which is characterised by the parametrical conditional probability distribution:

$$p_{\theta}(y|x) \equiv p(y|x, \theta) \quad (3)$$

where symbol ' \equiv ' means 'equal by definition'.

In the special case of repeated observations, with independent variations, we have:

$$p_{\theta}(y|x) = \prod_{i=1}^N p_{\theta}(y_i|x) \quad (4)$$

Let us now consider the *restitution process*.

In the general case it is characterised by the transformation:

$$p(x|y) = \int_{\Theta} p_{\theta}(y|x) \left[\int_X p_{\theta}(y|x) dx \right]^{-1} p(\theta) d\theta = \int_{\Theta} p_{\theta}(x, \theta|y) d\theta \quad (5)$$

This formula is the result of joint application of Bayes Theorem and of the principle of total probability (see Ref. [1] for details).

It is worth noting that on the basis of this result, *restitution may be interpreted as the probabilistic inversion of the mapping provided by the observation process, marginal with respect to the measurand*.

Now, restitution induces a probability measure on X and leads to the assignment of a probability distribution to x , namely $p(x|y)$. As a result of the measurement process it is also possible to describe the measurand by a single value, *the measurement value*, given by:

$$\hat{x} = E_x(x|y) \quad (6)$$

where E_x is the expectation operator, referred to the (random) variable x .

Finally, the overall *measurement process* may be viewed as a *probabilistic mapping from the set of the values of the measurand X to the same set of the measurement values $\hat{X} = X$* . Such a mapping is characterised by the conditional probability density $p(\hat{x}|x)$ which may be implicitly expressed by [27]:

$$p(\hat{x}|x) d\hat{x} = \int_{\Delta D\hat{x}} p(y|x) dy \quad (7a)$$

where $\Delta D\hat{x}$ is the region in the space of y where:

$$\Delta D\hat{x} = \{y : \hat{x} < E_x(x|y) \leq \hat{x} + d\hat{x}\} \quad (7b)$$

The essential formulas of the model are summarised in Table 1.

The model may be further generalised to accomplish for vector measurements, including dynamic measurements, and measurements based on

parameter estimation [1]. Anyway such generalisation will not be considered in this paper.

Table 1. A summary of the key formulas of the probabilistic model.

<i>The Reference Probabilistic Model of the Measurement Process</i>	
<i>Process / Sub-process</i>	<i>Probabilistic representation</i>
<i>Observation (sub-process)</i>	$p_{\theta}(y x) \equiv p(y x, \theta)$
<i>Restitution (sub-process)</i>	$p(x y) = \int_{\Theta} p_{\theta}(y x) \left[\int_x p_{\theta}(y x) dx \right]^{-1} p(\theta) d\theta$
<i>Measurement (process)</i>	$p(\hat{x} x) d\hat{x} = \int_{AD\hat{x}} \left[\int_{\Theta} p_{\theta}(y x) p(\theta) d\theta \right] dy$ $\Delta D\hat{x} = \{y : \hat{x} < E_x(x y) \leq \hat{x} + d\hat{x}\}$
where: $x \in X$: the measurand; $y \in Y$: N -dimensional observation vector; $\theta \in \Theta$: K -dimensional parameter vector; $\hat{x} \in X$: the measurement value.	

In the following sections some important models will be derived, as special cases of the above general model.

2.2 The mainstream-GUM model

The GUM [13] on the one hand presents general principles for the evaluation of measurement uncertainty, which are, as such, fairly independent from the assumed model, on the other hand develops a complete evaluation method, called in Ref. [6] “mainstream GUM”. We will also use this lexicon and also call “mainstream GUM model” the model implied in such a method. We want now to show its connection with our reference probabilistic model.

Let us then consider the case the observable output is given, for each value of the measurand x , by the following random function:

$$y = F(x) = f(x, \mathbf{v}) \quad (8)$$

where F is a random function, f is a deterministic function, \mathbf{v} is a vector of n random variables, the first m mutually correlated, the remaining $(n-m)$ uncorrelated. These random variables in general give rise to both observable (random) variations and to unobservable (systematic) deviations.

Let us assume f to be linear or locally linearisable so that, at least approximately and apart from an inessential additive constant, we have:

$$y \cong \kappa x + \mathbf{a}^T \mathbf{v} \quad (9)$$

where \mathbf{a} is a column vector of known sensitivity coefficients (and T is the transposition operator). Equation (9) may be trivially solved as:

$$\kappa \equiv x - \hat{x} \cong -c \cdot \mathbf{a}^T \mathbf{v} \quad (10)$$

where \hat{x} is the measurement value, κ is the correction to be (virtually) applied

to it to obtain the measurement value x , $c = k^{-1}$, $\hat{x} = cy$ or, in the case of repeated observations, $\hat{x} = c\bar{y}$.

Moreover, considering that y is a special case of \bar{y} and assuming the distributions of ν to be defined up to a vector of dispersion parameters δ , we get:

$$p(x|\bar{y}, \delta) = p_\kappa(x - \hat{x}|\delta) \Big|_{\hat{x}=\bar{y}} \quad (11)$$

which finally yields:

$$p(x|\bar{y}) = \int_{\delta} p(x|\bar{y}, \delta) p(\delta) d\delta \quad (12)$$

Equations (12) and (10) provide a probabilistic expression of the GUM Mainstream model. Additional details will be given in section 3.2, where numerical implementation is treated. We may notice that, in this model, the possibility of a deterministic inversion of the equation(s) defining the observation sub-process has been assumed, *which is not the general case*.

2.3 A stochastic observation scheme

The most general model at present implemented in the package (model UNCERT D, see section 3.3) is the following one:

$$\begin{aligned} y_i &= Q(kx + w_i + s + h) \\ s &= a^T \nu \end{aligned} \quad (13)$$

where:

$Q(\cdot)$ = quantisation operator, with quantisation interval q ;

w_i = stationary series of uncorrelated normal random variables, with variance σ^2 ;

s = influence quantity resulting from the linear combination of random variables, as in the previous model;

h = hysteresis term, characterised by $p(h) = \frac{1}{2}\delta(h - h_0) + \frac{1}{2}\delta(h + h_0)$,

where δ is here the Dirac-delta operator.

It is worth noting that, in this case, *no deterministic inversion (solution) of the model is possible*, due to the non-linear character of the quantisation operator, yet *the probabilistic inversion is still possible*.

The observation process is characterised by:

$$P(y|x, \sigma, s) = \frac{1}{2} \prod_i \int_{-q/2}^{+q/2} [p_w(y_i - x - s - h_0 + \xi|\sigma) + p_w(y_i - x - s + h_0 + \xi|\sigma)] d\xi \quad (14)$$

Here the dependence on the vector of dispersion parameters δ is implicit in the terms. Otherwise it may be made explicit as:

$$P(y|x, \sigma, \delta) = \int_S P(y|x, \sigma, s) p(s|\delta) ds \quad (15)$$

Restitution is then provided by:

$$p(x, \sigma|y) \propto \int_{\Delta} P(y|x, \sigma, \delta) p(\delta) d\delta \quad (16)$$

Finally the required distribution may be obtained as the marginal distribution:

$$p(x|y) = \int_{\Sigma} p(x, \sigma|y) d\sigma \quad (17)$$

Moreover, the distribution of σ may also be obtained:

$$p(\sigma|y) = \int_{\Sigma} p(x, \sigma|y) dx \quad (18)$$

Again, details on the implementation will be given later on, in section 3.3.

3 The Package UNCERT

3.1 General features of the package

The package has been developed with the aim of providing the metrologist with a set of tools giving full support in the expression of measurement results as a probability distribution over a suitable set of possible values of the measurand. Since direct calculation of involved distributions is performed, it is essential to consider an appropriate set of models to be treated. So the package is arranged in a hierarchical structure, starting from a set of basic modules, which may be extensively reused in the subsequent procedures, and progressively moving towards structures of increasing complexity. This philosophy has allowed a *highly modular design*. Parallel to the development of the programmes supporting such models, an implementation of the mainstream-GUM method has been done. This has a value in itself, but in the general economy of the package it is mainly intended to provide a *back-to-back testing* utility. Table 2 presents the progressive development of the package. This step-by-step approach has greatly favoured the validation process.

Another key implementation issue has been the choice of treating all involved values as *discrete random variables*. This choice has been motivated both on theoretical considerations and on implementation reasons, and related quantisation and truncation effects have been carefully studied and kept under control. Moreover, direct calculation of probability distributions essentially involves convolutions, in the case of linear relations, and weighted means among distributions, when non-linear, parametric dependency is to be accounted for.

Table 2. Steps in the development of package UNCERT

UNCERT 0-2: the GUM mainstream method			UNCERT A-D: Calculation of probability distributions	
Step	Level	Models	Level	Models
1	0	Infinite <i>dof</i> , uncorrelated variables	A	Infinite <i>dof</i> , uncorrelated variables
2	1	Finite <i>dof</i> , uncorrelated variables	B	Finite <i>dof</i> , uncorrelated variables
3	2	Finite <i>dof</i> , correlated variables	C	Finite <i>dof</i> , correlated variables
4			D	Stochastic observation scheme

In the following sections details will be presented on the implementation of the programmes forming the package.

3.2 Levels A through C

Levels A, B and C, progressively developed, allowed the expression of the final result of a measurement by a probability distributions, under the same hypothesis assumed in the GUM Mainstream model, presented in section 2.2.

In that model, it is possible to express the effect of all influence quantities on the final measurement result as a linear function of the variations in the quantities, i. e., neglecting an additive constant, as $e = c \sum_{i=1}^n a_i v_i$ or, equivalently, as $\kappa = -e$, with known coefficients c and a_i . For each of the input quantities, the following information is assumed to be available:

- either an estimated standard deviation $\hat{\sigma}_i$, with ν_i degrees of freedom, which may be finite or infinite;
- or an estimate of the range of variability², say $[v_{oi} - \Delta v_i, v_{oi} + \Delta v_i]$, with limit either certain or subject to a relative uncertainty, say $\alpha_i = \Delta \Delta v_i / \Delta v_i$;

Moreover, it is possible to identify a subset of mutually correlated input quantities, say, without loss of generality, $\mathbf{v}' = [v_1, v_2, \dots, v_m]$, $0 \leq m \leq n$, for which the corresponding correlation matrix is known $R = [r_{ij}]$.

Under the assumed hypotheses, we have: $p(x|\hat{x}) = p_\kappa(x - \hat{x})$ where $p_\kappa(x)$ may be expressed as the composition, through a convolution rule, of the

² We do not consider here, for simplicity, the case of a triangular distribution, which may however be treated in essentially a similar way as case b.

distribution accounting for the global influence of the subset of uncorrelated input quantities (if not null), $p_{\kappa, \text{uncorr}}(x)$, with the distribution accounting for the effect of complementary subset of mutually correlated input quantities (if not null), $p_{\kappa, \text{corr}}(x)$. Consequently:

$$p_{\kappa}(x) = p_{\kappa, \text{uncorr}}(x) * p_{\kappa, \text{corr}}(x) = \int p_{\kappa, \text{uncorr}}(x - \xi) p_{\kappa, \text{corr}}(\xi) d\xi \quad (19)$$

and we have to evaluate separately these two probability distributions.

For evaluating $p_{\kappa, \text{uncorr}}(x)$, it is necessary to assign a probability distribution to each of the variables, which may be done according to the rules summarised in Table 3.

Table 3. Calculation of standard deviations and probability distributions for individual influence quantities

Type of quantity	Dispersion parameter	Uncertainty on dispersion parameter	Resulting standard deviation	Probability distribution
<i>Normal</i>	Standard deviation σ_i	Infinite <i>dof</i>	$c a_i \sigma_i$	Normal
<i>Normal</i>	Standard deviation σ_i	Finite <i>dof</i>	$c a_i \sigma_i \nu_i / (\nu_i - 2)$	t Student
<i>Uniform</i>	Semi-range Δv_i	No uncertainty	$c a_i \Delta v_i / \sqrt{3}$	Uniform
<i>Uniform</i>	Semi-range Δv_i	Relative uncertainty on semi-range α_i	$\sigma_i = \sqrt{\frac{1}{3} \left(1 + \frac{\alpha_i^2}{3} \right)} \Delta v_i$	Logarithmic [28]

For the calculation of $p_{\kappa, \text{corr}}(x)$ we consider a normal distribution resulting from the sum of n-m variables, with correlation matrix equal to R , and standard deviations to be calculated according to the rules in Table 3.

3.3 Level D

This level corresponds to the implementation of the stochastic observation scheme, described in section 2.3.

The first implementation step requires calculation of $P(y|x, \sigma, s)$. Assuming a normal distribution for w , and defining the cumulative normal distribution as $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\xi^2/2} d\xi$ and considering a quantisation of variables x, σ, s , with quantisation interval Δx , for each observed value y_i , we have:

$$P(y_i | x, \sigma, s) = \frac{1}{2} \Phi \left(\frac{y_i + \frac{q}{2} - kx - s - h_0}{\sigma} \right) - \frac{1}{2} \Phi \left(\frac{y_i - \frac{q}{2} - kx - s - h_0}{\sigma} \right) + \frac{1}{2} \Phi \left(\frac{y_i + \frac{q}{2} - kx - s + h_0}{\sigma} \right) - \frac{1}{2} \Phi \left(\frac{y_i - \frac{q}{2} - kx - s + h_0}{\sigma} \right) \quad (20)$$

For the whole vector of observations y , we obtain:

$$P(y | x, \sigma, s) = \prod_i P(y_i | x, \sigma, s) \quad (21)$$

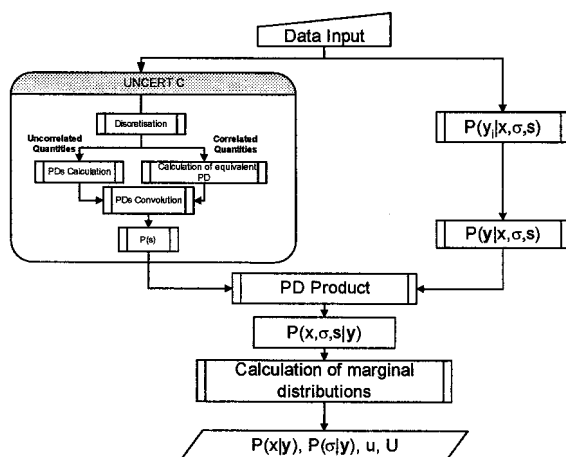


Figure 1. Flowchart of UNCERT D

Now it is possible to calculate the joint distribution $P(x, \sigma, s | y)$:

$$P(x, \sigma, s | y) = P(y | x, \sigma, s) P(x) P(\sigma) P(s) \left[\sum_{x, \sigma, s} P(y | x, \sigma, s) P(x) P(\sigma) P(s) \right]^{-1} \quad (22)$$

where:

- $P(x)$ is assumed to be constant over the set of the possible values of x ;
- $P(\sigma)$ is either supplied by the user or assumed to be constant over the set of the possible values of σ ;
- $P(s)$ is either supplied as output of UNCERT C Programme or assumed to be constant over the set of the possible values of s .

Finally the following marginal distributions may be calculated:

$$P(x|y) = \sum_{\sigma, s} P(x, \sigma, s|y) \quad (23a)$$

$$P(\sigma|y) = \sum_{x, s} P(x, \sigma, s|y) \quad (23b)$$

The flowchart of the algorithm is presented in Fig 1.

3.4 Software quality assurance

The quality assurance of software intended for use in metrology is a very important issue, subject of specific research effort [18-21]. It includes general quality requirements on the quality of software [22-23] as well as more specific metrological requirements [20-21]. Moreover, the package UNCERT has been developed in a *research environment*, which adds an additional important feature, since quality in research laboratories is an open problem, due to the specific character of research activities [24-26]. Let then us briefly address the following key points:

- the development environment;
- the implementation process;
- the validation process.

The development of the package has been realised in a research context, in parallel with the development of the related theory, mainly concerning the probabilistic modelling of measurement processes, and with studies on the evaluation of uncertainty in industrial environment, supported by dedicated experimentation. In the first stage, software has been intended mainly to support, through numerical examples, the theoretical studies. In a second stage, while the theory reached a first satisfactory systematization, a methodical implementation has been undertaken, according to quality assurance principles and the metrological requirements. The goal of this second phase is to propose the package as a *standard tool* for the expression of measurement results in probabilistic terms, by direct calculation of related probability distributions. The personnel involved in the development of the package have a strong scientific and metrological background and good programming experience. The development of the software has been done according to internal quality assurance procedure of the laboratory, which is undergoing ISO 9001 certification.

The implementation process has included the definition of the overall architecture of the package, based on the hierarchy of models detailed in section 3.1, and the implementation of the individual modules, corresponding to the “levels” in Table 2.

The main steps in the implementation of the individual modules, according to the laboratory procedure for the development of proprietary software, has been the following:

1. definition of the specification of the module;
2. preparation of the code;
3. planning of a set of tests (*validation suite*);
4. application of the tests and eventual correction/improvement of the module;
5. approval of the module and its formal inclusion in the software databases of the laboratory, on a dedicated area, located in the laboratory-internal-network server; from this moment onwards, all successive modifications must be approved and documented;
6. each approved module is accompanied by:
 - the *manual*, which is at the same time a user's manual, containing all necessary information for the correct use of the module, and a programmer's manual, containing all the necessary details on the implementation of the module;
 - a detailed report on the validation campaign;
 - the input and output data set of all the validation tests performed; a part of such tests are intended to be repeated, each time a variation is applied, as a part of the variation-approval procedure;
 - occasionally, dedicated studies on special features of the module have been performed and reported; for example a careful study of the effects of quantisation and truncation has been produced.

Finally, the validation procedures is based on the following main principles:

1. since the programs works on discrete variables, distributions with a very small number of points (for example three points), have been used to check the functionality, by independent calculation of the expected results;
2. on the opposite side, *stress testing* has been applied, concerning for example checking performance for large number of points per distribution (order of 10^3) and large number of variables (10^1);
3. *back-to-back comparison* with the programmes implementing the GUM-Mainstream method has been employed;
4. reference data have been used, when possible, taken typically from the GUM or from other standards.

Further validation is expected to take place, through the network SoftTools_MetroNet.

4 Application Examples

4.1 Calibration of an end-gauge

Let us consider example H1 of the GUM, which concerns the measurement of standard reference blocks. Uncertainty evaluation is carried out after linearization of the measurement model, according to the GUM mainstream procedure, and Table 4 presents the input data.

Table 4. Summary of the input data for the GUM H1 example.

Quantity	Description	Probability Distribution	Dispersion parameter	Uncertainty on dispersion parameter
1	Calibration	t-Student	$\sigma = 25$	18 <i>dof</i>
2	Measured difference	t-Student	$\sigma = 9.7$	26 <i>dof</i>
3	Expansion coefficients	Logarithmic	$\Delta v = 2.9$	10% range
4	Temperatures of gauges	Logarithmic	$\Delta v = 16.6$	50% range

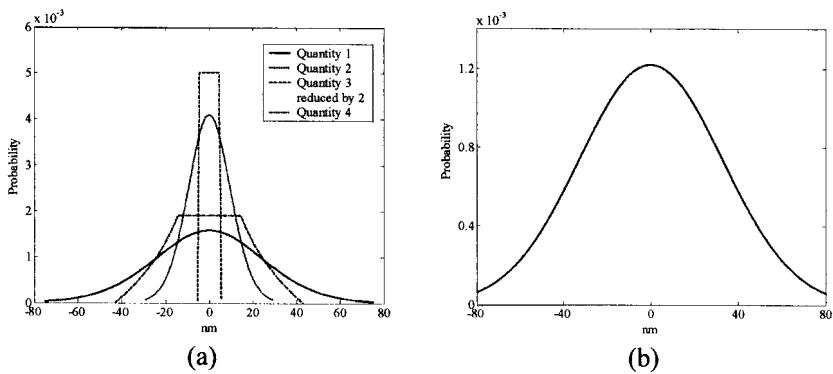


Figure 2 Test case 1: probability distributions of the input quantities (a) and final distribution on a separate scale (b).

4.2 Repeated observations

Let us consider an example concerning repeated observations, from a low-resolution indicating device (adapted from [29]), considering two cases:

- A) the instrument is free from systematic deviations
- B) the instrument is assumed to have systematic deviations, bounded within $\pm 0.05 \text{ mm}$. Numerical values are collected in Table 5.

Table 5 Numerical data for cases study. All data are expressed in millimetres.

Test case	q	Δs	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}
A	0.1	0	7.5	7.5	7.4	7.5	7.5	7.4	7.5	7.5	7.5	7.4
B	0.1	0.05	7.5	7.5	7.4	7.5	7.5	7.4	7.5	7.5	7.5	7.4

Since observations have a resolution limited to q , we will discretise all relevant quantities with $\Delta x = q/10$.

The resulting discrete joint distribution $P(x_i, \sigma_k | y)$ is presented in Fig. 3.

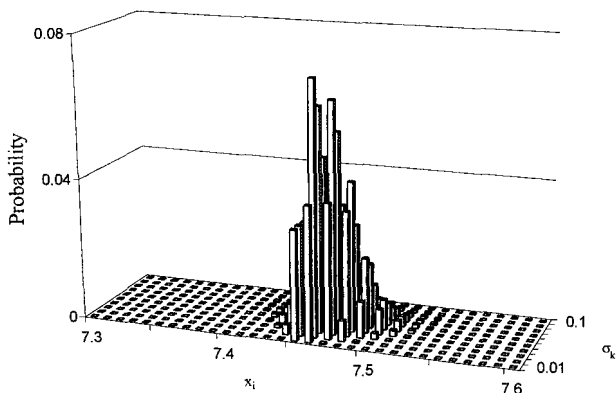


Figure 3 The graph of $P(x_i, \sigma_k | y)$

The final measurement results, i. e. the density to be assigned to x and to σ are shown in Figure 4.

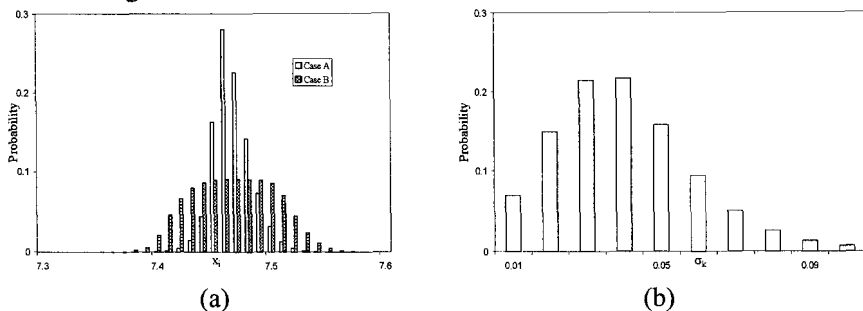


Figure 4. (a) Final result of measurement in both cases A and B. (b) Graph of $P(\sigma_k | y)$.

4.3 Measuring device with hysteresis

Let us consider now a case where some hysteresis is present in the measuring device. We consider two set of (simulated) data: in the first set the instrument is assumed to work on the $+h_0$ curve, in the second to randomly commute between the two curves. Results are presented in Fig.5. Now it is interesting to note that the probabilistic restitution algorithm is able to automatically recognise the two situations and give the correct results. In particular it emerges how the second case is more favourable, since data are unbiased, although there is a greater dispersion.

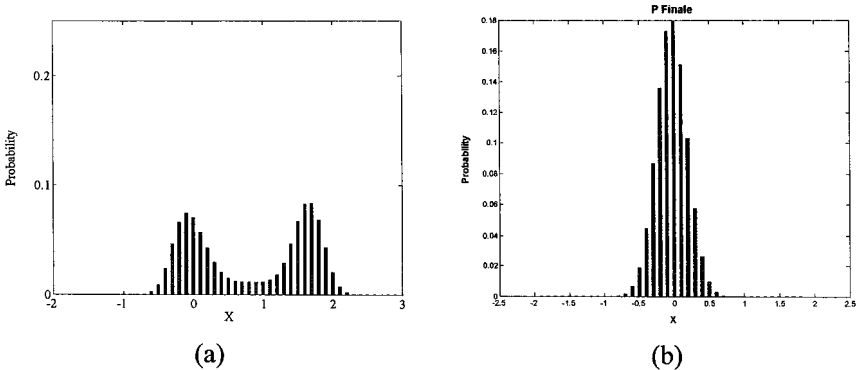


Figure 5. Results from a measuring devices affected by hysteresis.

5 Conclusions

In this paper the possibility of expressing the final result of a (any) measurement by a probability distribution over the set, discrete or continuous, of the possible values of the measurand has been considered.

In particular, a reference probabilistic model has been presented showing how it includes, as a special case, the model assumed in the GUM Mainstream method and also allows treatment of more sophisticated models, based on a stochastic observation scheme, including non linearities such as quantisation and hysteresis. Then the related software tools have been addressed, with reference to the package UNCERT, which has been developed with the aim of providing the metrologist with a set of tools giving full support in the expression of measurement results as a probability distribution. The package is organised on the basis of a hierarchy of models, which allows a highly modular design, and is based on direct calculation of involved probability distributions, treating all variable as discrete. It has been developed according to quality assurance criteria and validated by wide testing and is intended to be proposed as a *standard tool* for this kind of calculations. The expression of measurement results in probabilistic terms may be particularly beneficial whenever risk assessment or cost evaluation is required.

6 Acknowledgments

The package UNCERT has been developed within the network SofTools MetroNet. The research has been partially funded within a National Research Programme.

References

1. G. B. Rossi, *Measurement* **34**, 85-99 (2003)
2. A. Hetman, M.J. Korczynski, *Proc. 2nd on line IMEKO TC 1* (2002)
3. B. D. Hall, *Measurement Science and Technology* **10**, 380-385 (1999)
4. P. Otomanski, *Proc. IMEKO TC7 Symposium* (2002)
5. Spec. issue "The foundation of measurement", *Measurement* **34** (1) (2003)
6. M.G. Cox, P.M. Harris, *Measurement of uncertainty and the propagations of distributions*, NPL (2002)
7. www.bipm.fr/enus/2_Committees/JCGM.shtml.
8. INERIS European Conf. "Environment, health, safety, a challenge for measurement (2001)
9. W. Schulz, K.D. Sommer, *OIML Bulletin* **XL** 4, 5-15 (1999)
10. G.B. Rossi, F. Crenna, *Proc. IMEKO TC7 Symposium* (2002)
11. G. B. Rossi, F. Crenna, *SIMAI Conference* (2002)
12. BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *International Vocabulary of Basic and general terms in Metrology*, Second Edition (1994)
13. BIPM,IEC,IFCC,ISO,IUPAC,IUPAP,OIML, *Guide to the expression of uncertainty in measurement*, ISO Genève (1993)
14. Spec. issue on "Meas. and instrum. science", *Measurement* **14** (1) (1994)
15. M. G. Cox, in: *Advanced mathematical tools in Metrology*, 1-22, World Scientific (1994)
16. M.G. Cox, P.M. Harris, M.P. Dainton, *Uncertainty and statistical modelling-Software support for metrology*, NPL (2001)
17. K. D. Sommer, M. Kochsiek, B. Siebert, A. Weckenmann, *Proc IMEKO XVII* 1248-1253 (2203)
18. D. Richter, in: *Advanced mathematical tools in Metrology*, 255-262, World Scientific (1994)
19. M.G. Cox, P.M. Harris, M.P. Dainton, *Software specification for uncertainty calculation and associated statistical analysis*, NPL (2001)
20. B. Wichmann, *Validation in measurement software*, NPL (2001)
21. D. Brinkley, *Guidance on developing software for metrology*, NPL (2001)
22. L. Bernstein, C. M. Yuhas, in: *Juran's Quality Handbook*, 20.1-22, McGraw-Hill (2000)
23. G. Satriani, in: *Manuale della Qualità*, 583-614, Il Sole 24 ORE (2001)
24. G. B. Rossi, P. Salieri, S. Sartori, *Measurement*, **32** 117-123 (2002)
25. Al C. Endres, in: *Juran's Quality Handbook*, 19.1-38, McGraw-Hill (2000)
26. G. B. Rossi, F. Crenna, F. Pampagnin, *Proc. Euspen* 1, 439-442 (1999)
27. A. Papoulis, *Probability random variables and stochastic processes*, McGraw Hill, (1984)
28. I. Lira, *Evaluation the measurement uncertainty*, IOP (2002)
29. I. Lira, W. Woger, *Meas. Sci. Technol.*, **9** 1167-73 (1998)
30. E.O. Doebelin, *Measurement Systems*, McGraw Hill, (1990)

FEASIBILITY STUDY OF USING BOOTSTRAP TO COMPUTE THE UNCERTAINTY CONTRIBUTION FROM FEW REPEATED MEASUREMENTS*

B. R. L. SIEBERT

*Physikalisch-Technische Bundesanstalt
Bundesallee 100, D-38116 Braunschweig, Germany
E-mail: Bernd.Siebert@ptb.de*

P. CIARLINI

*CNR, Istituto per le Applicazioni del Calcolo, "M. Picone"
V.le del Policlinico 137, 00161 Roma, Italy
E-mail: ciarlini@iac.rm.cnr.it*

Information on possible values of a quantity can be gained in various ways and be expressed by a probability density function (PDF) for this quantity. Its expectation value is then taken as the best estimate of the value and its standard deviation as the uncertainty associated with that value. A coverage interval can also be computed from that PDF. Information given by a *small* number n of values obtained from repeated measurements requires special treatment. The Guide to the Expression of Uncertainty in Measurement recommends in this case the t -distribution approach that is justified if one knows that the PDF for the measured quantity is a Gaussian. The bootstrap approach could be an alternative. It does not require any information on the PDF and -based on the plug-in principle- can be used to estimate the reliability of any estimator. This paper studies the feasibility of the bootstrap approach for a *small* number of repeated measurements. Emphasis is laid on methods for a systematic comparison of t -distribution and bootstrap approach. To support this comparison, a fast algorithm has been developed for computing the total bootstrap and total median.

1. Introduction

The Guide to the Expression of Uncertainty in Measurement [1], briefly termed the GUM, distinguishes for historical reasons between the so-called Type A and Type B evaluation of the standard uncertainty. The first treats information gained from repeated measurements and the latter any other information. Both types of evaluation lead to a probability density function (PDF) for the considered input quantity X . The expectation value x of that PDF is taken as best estimate for the value of X and its standard deviation as the uncertainty $u(x)$ associ-

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

ated with that estimate. The PDF for the output quantity Y can be computed if one knows the PDFs for the input quantities and the model that relates them to the output quantity. The best estimate y and the uncertainty $u(y)$ are related to the PDF for Y in the same manner as described for input quantities.

The functional form (shape) of the PDF for the output quantity Y needs to be known for determining the so-called expanded uncertainty U_p . The latter is used to indicate an interval about the best estimate y that - *based on the information given* - contains a large fraction p of the distribution of values that could reasonably be attributed to the output quantity Y .

The standard GUM procedure for evaluating uncertainty is valid for linear models. In this case, the PDF for the output quantity is normal (Gaussian) unless one encounters one or few dominating input quantities with non-normal PDFs. Frequently occurring examples are a rectangular PDF or contributions from quantities assessed by only a *small* number n of repeated measurements.

The standard procedure of the GUM uses the *sample mean* as best estimate x for the value of X and the *experimental standard deviation of the mean* as uncertainty associated with x and infers the expanded uncertainty from a t -distribution with $\nu=n-1$ degrees of freedom. This approach is justified if the PDF for the measured input quantity is a Gaussian.

The bootstrap approach may provide an alternative to the standard Type A evaluation of uncertainty. The bootstrap is known to give reliable estimates of estimators and their uncertainty for large sample sizes for any PDF and it has been used in metrological applications since many years [2].

This paper examines the feasibility of the bootstrap approach for *small* sample sizes. It describes briefly the t -distribution and bootstrap approach, provides computational tools for fast bootstrapping that allow the total bootstrap and median to be enumerated. It compares systematically both approaches to Type A evaluation by taking many samples of a given small size from PDFs encountered in practice. Finally, it presents conclusions drawn from the comparison.

2. Approaches for Type A evaluations

2.1. The t -distribution approach

The GUM [1] uses an appropriate t -distribution for the quantity X if the information on possible values of a X consists of a set of n values $\{x_1, \dots, x_n\}$ obtained from repeated measurements under repeatability conditions.

In this approach, the *mean value* x of the measured values is the best estimate for the value of the quantity X and the *experimental standard deviation of the mean* is the uncertainty $u(x)$ associated with that mean from the set of n values $\{x_1, \dots, x_n\}$:

$$x = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{and} \quad u^2(x) = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - x)^2. \quad (1)$$

Based on Bayesian probability theory [3], one obtains a t -distribution with $\nu = n-1$ degrees of freedom for the quantity X , if one has the additional information that the measured values x_i are drawn from a Gaussian PDF. The t -distribution with ν degrees of freedom and the transformation to the PDF for X is given by:

$$g_{T,\nu}(\tau) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma(\nu/2)} \left[1 + \frac{\tau^2}{\nu}\right]^{-\frac{\nu+1}{2}} \quad \text{and} \quad \tau = \frac{\xi - x}{u(x)}, \quad (2)$$

where τ and ξ denote possible values. The PDF for X is analogously expressed by $g_X(\xi)$ and also denoted by X^c .

The t -distribution is symmetric and the boundaries of the coverage interval for a coverage probability p , can be derived from

$$p = \int_{-t_{p,\nu}}^{t_{p,\nu}} g_{T,\nu}(\tau) d\tau \quad \text{and} \quad U_p \equiv [x - u(x)t_{p,\nu}, x + u(x)t_{p,\nu}]. \quad (3)$$

We use $p=0.95$ in this paper and set $x - u(x)t_{p,\nu} = t_{0.025}$ and $x + u(x)t_{p,\nu} = t_{0.975}$.

2.2. The Bootstrap approach

The nomenclature used in this paper attempts to follow the GUM. This implies that the bootstrap nomenclature is to be adjusted in some details.

In practice, the PDF that can be reasonably be assumed to underlie the dispersion of the measurements is not necessarily normal. The bootstrap approach does not require any information on the PDF from which the measured values x_i are drawn. It was designed to estimate the reliability of any estimator [4]. It is justified by the *plug-in principle* that is visualised in Figure 1. This principle can be adopted for any statistics, not only for the mean. For instance, with a simple formula it is possible to compute the estimate of the variance of the median. The principle states that the *bootstrap replications* of any statistics applied to *bootstrap samples* mimic the statistics of the corresponding output PDF. A bootstrap

^c Latin letters denote quantities (upper case) and the best estimates of their value (lower case). Lower case Greek letters indicate possible values.

sample is generated by *re-sampling with replacement* from the measured sample (data set) of n values $\{x_1,..., x_n\}$. We denote the bootstrap sample by $\{\xi_i,...\xi_n\}_r$; the subscript r runs from 1 to R . For each of the R bootstrap samples one computes a bootstrap replication of the statistics of interest, *i.e.* in this paper the mean value x_r . The bootstrap replications of x_r mimic the “PDF”^d for the mean. We calculate the arguments $b_{0.025}$ and $b_{0.975}$ at which the cumulative “PDF”, termed CDF, of the mean reaches the probability indicated by the indices. These values give the endpoints of a probabilistically symmetric 95 % coverage interval, but not in general the shortest

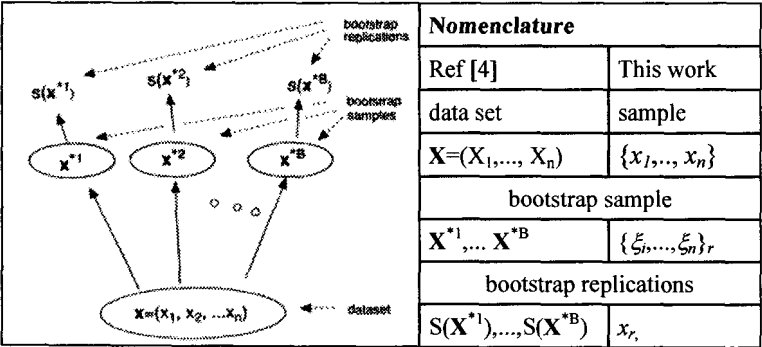


Figure 1. Visualisation of the plug-in principle [4] and nomenclature used in this paper. The bootstrap samples are usually numbered from 1 to B but from 1 to R in this work

Figure 2 illustrates and compares the t -distribution and bootstrap approach to infer the distribution of the sample mean for a sample of 5 and 8 simulated *measured* values. The t -distribution has been computed using Equation (3). The bootstrap will be explained in detail below in Section 3.

^d The notation “PDF” indicates that it is mimicked. A mimicked bootstrap PDF is not continuous but a set of at most $(2n-1)!/n!(n-1)!$ discrete increasingly ordered values.

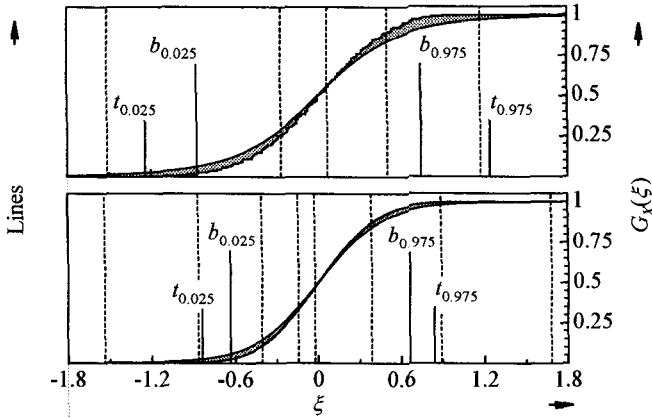


Figure 2. Bootstrap and t -distribution approach for the sample mean of samples $\{x_1, \dots, x_5\}$ (upper part) and $\{x_1, \dots, x_8\}$ (lower part). The values x_i (drawn from a Gaussian PDF) are indicated by broken lines. The smooth curves are the cumulative t - and total bootstrap-distributions (CDF) of the sample mean. For $t_{0.025}$, $t_{0.975}$, $b_{0.025}$ and $b_{0.975}$ see the main text.

3. Computational tools for a systematic comparison

To study the feasibility of the bootstrap approach, it is necessary to examine a large number M of samples in order to obtain a representative data basis. In order to demonstrate that the bootstrap works for any distribution it is necessary to consider several PDFs that might occur in practice. In this paper we consider a Gaussian (G), rectangular (R), triangular (T) and U-shaped PDF (U) and use the generic term *parent*-PDF for these PDFs. Furthermore, we study sample sizes ranging from 4 to 8 and compute for each sample the expanded uncertainty based on the t -distribution and the bootstrap approach.

To determine the expanded uncertainty using the t -distribution approach we compute x_m and $u(x_m)$ for $\{x_1, \dots, x_n\}_m$ from Equation (1); m runs from 1 to M and indicates the sample (simulated set of measured data). Equation (3) then delivers $t_{0.025}$ and $t_{0.975}$. The values of $t_{p,3}$ to $t_{p,7}$ are 3.182, 2.776, 2.571, 2.447 and 2.365.

The calculation of expanded uncertainty using the bootstrap approach requires more effort. For each sample of measured values $\{x_1, \dots, x_n\}$ one needs to calculate R bootstrap samples $\{\xi_i, n\}_r$ and generate bootstrap replications of the mean $x_{m,r}$. The bootstrap replications mimic the “PDF” of the mean from which we calculate $b_{0.025}$ and $b_{0.975}$. $t_{0.025}$, $t_{0.975}$, $b_{0.025}$ and $b_{0.975}$ have the same physical dimension as the quantity X .

Without loss in generality we use the parent-PDFs for X in standard form, i.e. $x=0$ and $u(x)=1$. The *objective of the study* is then reduced to examine whether the intervals $[t_{0.025}, t_{0.975}]_m$ and $[b_{0.025}, b_{0.975}]_m$ encompass the value zero. We denote the number of samples for those $t_{0.025} > 0$ by M_{t-} and define the ratio of M_{t-} to M as the probability p_{t-} for violating the “ t -criterion”; p_{t+} , p_{b-} and p_{b+} are defined analogously.

The number M of simulated sets of measured data and the number R of bootstrap replications that are needed to achieve representative results depend on the method of calculation and the sample size n .

3.1. “Brute force” and total bootstrap

Let us first consider a “brute-force” approach. In that case one uses Monte Carlo to generate M samples and again Monte Carlo to generate R bootstrap replications for each sample. The trial numbers M and R would be determined by the usual convergence criteria; however usually $M \geq 10^6$ and $R \geq 10^3$ [4].

This “brute-force” approach quite long computing times for the calculations themselves and the convergence study. For small sample sizes n one achieves a considerable reduction of computing time by using the *total bootstrap* instead of Monte Carlo. To some extent it is also possible to replace the Monte Carlo simulation of the samples (sets of measured data) by an considerably faster method that will be described briefly in Section 3.3.

The total number of bootstrap samples and replications R for a given sample of size n is equal to n^n (sampling with replacement!), e.g. for $n=8$ one obtains $R=16\,777\,216$. However, many of these bootstrap samples consist of the same values. Consider for example the bootstrap sample $\{\xi_1=x_1, \dots, \xi_8=x_8; 8\}_r$; there are $40\,320$ (i.e. $n!$) permutations that yield the same value for x_r . The idea behind the total bootstrap is quite simple: one uses only *one* bootstrap sample and weights its contribution according to the number of possible permutations. In the example above one obtains as weight $n!/n^n$. The reduction of R is considerable since $R_{\text{total}} = (2n-1)!/(n!(n-1)!)$ whence, for $n=8$, $R_{\text{total}} = 6\,435$. In our study it is quite important to represent the tails of the “PDF” of the bootstrap replications. These are exactly represented by the total bootstrap but, since the tails have very small weights, a large R would be needed when using Monte Carlo for the generation of bootstrap samples.

To enumerate the total bootstrap, we consider the values x_i of the sample $\{x_1, \dots, x_n\}_m$ as n “peas” with n different colours, simply denoted by $1, 2, \dots, n$, and assume an urn that contains an equal but indefinitely large numbers of peas of each colour. A bootstrap sample is then realized by drawing n peas from that

urn. We denote a possible set of coloured peas by $[i_1, i_2, \dots, i_n]$ and we want to form all sequences of n digits drawn from $\{1, \dots, n\}$ that are non-decreasing, e.g., when $n = 3$, the sequences are 111, 112, 113, 122, 123, 133, 222, 223, 233, 333. The process is one of *counting*, from $1 \dots 1$ to $n \dots n$, but omitting all numbers containing digits that decrease. We could also say that we apply the multinomial theorem to count the permutations. Thus, with $n = 5$, there are $5!/(1!2!0!0!2!) = 30$ permutations of 12255. We use a simple algorithm that carries out the drawing of the R_{total} bootstrap samples of interest. The algorithm for the enumeration of the *total bootstrap* is demonstrated in Table 1 and consists of the following steps:

- ① Start with $[1, 1, \dots, 1]$, i.e. in all positions are peas of colour 1.
- ② Increment by one the colour number in the rightmost position $j_{\text{②}}$ that contains a color number of less than n . Repeat that step until the colour n in position $j_{\text{②}}$ is reached, and then proceed to step ③.
- ③ Find the rightmost position $j_{\text{③}}$ that contains a colour number of less than n and increment by one the colour number in position $j_{\text{③}}$ by 1 and set the colour number in position $j_{\text{③}}$ equal to that in position $j_{\text{②}}$. If possible, proceed with step ②; otherwise proceed with step ④.
- ④ Stop if neither step ② nor step ③ is possible $j_{\text{④}}$, i.e. when $[n, n, \dots, n]$ is reached.

Table 1. Algorithm for the generation of the *total bootstrap* and the *total median*. A sample of size $n=4$ is used. For lack of space, only a subset of the R entries, ordinarily numbered by r , is shown. The number of permutations m_r divided by n^n is the probability occurrence. Column C displays the steps (see main text) that lead to entry shown.

r	m_r	1	2	3	4	C
1	1	1	1	1	1	①
2	4	1	1	1	2	②
3	4	1	1	1	3	②
4	4	1	1	1	4	②
5	6	1	1	2	2	③

r	m_r	1	2	3	4	C
7	12	1	1	2	4	②
8	6	1	1	3	3	③
15	24	1	2	3	4	②
34	4	3	4	4	4	②
35	1	4	4	4	4	④

Consider as an example for step ② the step from $r=1$ to $r=2$, i.e. from $[1, 1, 1, 1]_1$ to $[1, 1, 1, 2]_2$. in Table 1: $j_{\text{②}}=n$, and the colour number in position $j_{\text{②}}$ is raised from 1 to 2. Consider as an example for step ③ the step from $r=4$ to $r=5$, i.e. from $[1, 1, 1, 4]_4$ to $[1, 1, 2, 2]_5$. in Table 1: $j_{\text{③}}$ the value 3 in step ③; therefore, the colour number in position $j_{\text{③}}$ is raised from 1 to 2 and the colour number in position $j_{\text{②}}$ is set to 2.

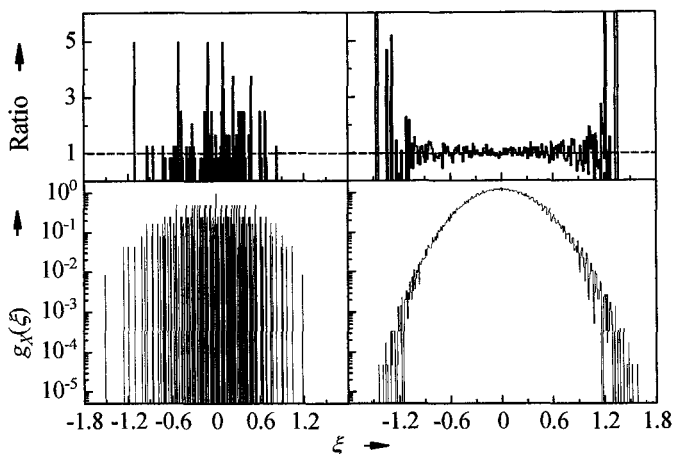


Figure 3. Comparison of total and “brute force” bootstrap. The left figures pertain to $n=5$ and the right ones to $n=8$. The figures show the “PDF” of the sample mean mimicked by the total bootstrap (bottom) and (top) the ratio of the “PDFs” mimicked by the total bootstrap to that mimicked by the “brute-force” bootstrap. The corresponding cumulative distributions were shown in Figure 2. The bootstrap gives always gives *discrete* values (see footnote d), but the resolution in graphical representations is too coarse to show that.

Figure 3 compares the total bootstrap for samples with $n=5$ and $n=8$; R_{total} is 126 and 6435, respectively. For the normal bootstrap we used $R = R_{\text{total}}$. The figure shows clearly that the total bootstrap gives a far better representation of the tails of the PDF for the mean value. Using Monte Carlo we computed for many samples the probabilities for violating the t -criterion and the b -criterion using the total and the normal bootstrap and produced frequency distributions of the corresponding ratios. The results showed that the total bootstrap is much faster and more reliable for the studied sample sizes.

3.2. Total bootstrap and the median

Following an idea of Cox and Pardo [5], we also calculate the weights of an increasingly ordered sample for the determination of the median and the uncertainty associated with it. We reproduce their result for a sample size $n=7$ and extend the calculation for sample sizes up to 17. The median of an ordered *odd* sample is the $((n+1)/2)^{\text{th}}$ element. The total median can be represented in this case by a simple vector of weights $p_{n,\text{odd}}$ with the elements $p_{i,n}$. The total median of an ordered *odd* sample is symmetric, *i.e.* $p_{i,n} = p_{n-i,n}$. The results for *odd* sample sizes are: $p_{1,3} = 0.25926$, $p_{2,3} = 0.48148$ and the results for *odd* sample sizes from $n=5$ up to $n=17$ are shown in Table 2.

Table 2. Weights of the total median for odd sample sizes. Since the total median is in this case symmetric only the first $(n+1)/2$ element are displayed.

i	$p_{i,7}$	$p_{i,9}$	$p_{i,11}$	$p_{i,13}$	$p_{i,15}$	$p_{i,17}$
1	0.01500	0.00145	0.00017	0.00002	0.00000	0.00000
2	0.09812	0.02892	0.00703	0.00146	0.00027	0.00004
3	0.23863	0.11447	0.04404	0.01423	0.00397	0.00098
4	0.30620	0.22066	0.12150	0.05495	0.02121	0.00716
5	-	0.26899	0.20588	0.12427	0.06278	0.02744
i	$p_{i,5}$	$i=6$	0.24274	0.19361	0.12487	0.06844
1	0.05792	$i=7$	-	0.22294	0.18324	0.12429
2	0.25952	$i=8$	-	-	0.20732	0.17435
3	0.36512	$i=9$	-	-	-	0.19458

Using the data provided in Table 2, one can compute the total median and the uncertainty associated with it for odd sample sizes from Equation (4):

$$x_{\text{median,odd}} = \sum_{i=1}^n p_{i,n} x_i \text{ and } u^2(x_{\text{median,odd}}) = \sum_{i=1}^n p_{i,n} (x_i - x_{\text{median,odd}})^2. \quad (4)$$

The median of an ordered *even* sample is the mean of the central pair, i.e. of the $(n/2)^{\text{th}}$ and $(n/2+1)^{\text{th}}$ element [5]. Therefore, to represent the results for *even* sample sizes one needs a Matrix $P_{n,\text{even}}$ with the elements $p_{i,j,n}$, where i is the $(n/2)^{\text{th}}$ and j the $(n/2+1)^{\text{th}}$ element of the increasingly ordered sample $\{x_1, \dots, x_n\}$. Obviously, the elements $p_{i,j,n}$ do not exist since the samples are ordered. The matrix is therefore an upper triangular matrix. It is symmetric about the anti-diagonal elements $p_{i,n+1-j,n}$, i.e., $p_{i,j,n} = p_{n+1-j,n+1-i,n}$.

The value of the median and the uncertainty associated with it can be computed from Equation (5a) and Equation (5b):

$$x_{\text{median,even}} = \sum_{i=1}^n \sum_{j=i}^n p_{i,j,n} \frac{1}{2} (x_i + x_j) \quad (5a)$$

$$u^2(x_{\text{median,even}}) = \sum_{i=1}^n \sum_{j=i}^n p_{i,j,n} \left[\frac{1}{2} (x_i + x_j) - x_{\text{median,even}} \right]^2. \quad (5b)$$

Equation (5a) can be re-ordered to

$$x_{\text{median,even}} = \sum_{i=1}^n \frac{1}{2} \left(\sum_{j=i}^n p_{i,j,n} + \sum_{k=1}^i p_{k,i,n} \right) x_i. \quad (5c)$$

The term in brackets in Equation (5c) is the sum over the i^{th} row and the i^{th} column of Matrix $P_{n,\text{even}}$. We denote half of this sum by $p'_{i,n}$ and replace Equation (5a) by Equation (6):

$$x_{\text{median,even}} = \sum_{i=1}^n p'_{i,n} x_i. \quad (6)$$

Table 3a shows the weights $p'_{i,n}$ for sample sizes from $n=4$ to $n=16$ and Table 3b shows the elements of the median Matrix $P_{8,\text{even}}$. Unfortunately, a corresponding rearrangement of Equation (5b) is not possible.

Table 3a. Weights of the total median for a sample sizes from $n=4$ to $n=16$ for use with Equation (6). The weights not shown can be inferred from the symmetry $p'_{i,n} = p'_{n+1-i,n}$.

i	$p'_{i,6}$	$p'_{i,8}$	$p'_{i,10}$	$p'_{i,12}$	$p'_{i,14}$	$p'_{i,16}$
1	0.03549	0.00624	0.00089	0.00010	0.00001	0.00000
2	0.17438	0.06432	0.01869	0.00450	0.00093	0.00017
3	0.29012	0.17246	0.07923	0.02972	0.00944	0.00261
4	-	0.25698	0.16776	0.08776	0.03845	0.01453
5	-	-	0.23343	0.16243	0.09287	0.04529
i	$p'_{i,4}$	$i=6$		0.21548	0.15712	0.09594
1	0.15625	$i=7$			0.20118	0.15205
2	0.34375	$i=8$				0.18942

Table 3b. The total median matrix P for a sample of size $n=8$.

i	$p_{i,3,8}$	$p_{i,4,8}$	$p_{i,5,8}$	$p_{i,6,8}$	$p_{i,7,8}$	$p_{i,8,8}$
1	0.00280	0.00154	0.000730	0.00027	0.00006	0.00000
2	0.04199	0.02309	0.010952	0.00407	0.00094	0.00006
3	0.06532	0.10007	0.04746	0.01763	0.00407	0.00027
4	-	0.10117	0.12778	0.04746	0.01095	0.00073
5	-	-	0.10117	0.10007	0.02309	0.00154
	$p_{1,1,8} = 0.001230$		$i=6$	0.06532	0.04199	0.00280
	$p_{1,2,8} = 0.004610$		$i=7$	-	0.02146	0.00461
	$p_{2,2,8} = 0.021457$		$i=8$	-	-	0.00123

Figure 4 shows for the samples already used in Figures 1-3 the CDFs for the total median using the data in Tables 3a and 3b and compares them with CDFs for the mean mimicked by the replications from the total bootstrap. As expected from theory, the median distribution is “broader” and reflects some “memory” of the sample values.

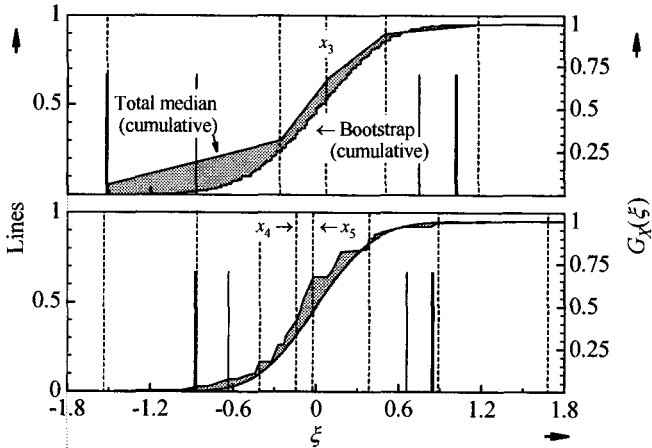


Figure 4. Comparison of the CDF for the total bootstrap replications of the mean (histogram) and the total median (step-wise linear) for a sample of size $n=5$ (top) and $n=8$ (bottom). The values x_i (same as in Figures 2 and 3) are indicated by broken lines. The thin solid lines indicate $b_{0.025}$ and $b_{0.975}$. The thick lines indicate the arguments where the cumulative median distribution reaches the values 0.025 and 0.975, respectively.

3.3. The bootstrap approach and discretely represented distributions

The benefit in using the total bootstrap algorithm has been a considerable reduction in the computing time. We found a very simple method to simply extend the ideas used for enumerating the total bootstrap. We consider n_{arg} positions that we interpret as measured values. Our urn contains now an equal but indefinitely large numbers of peas of each colour for any of the n_{arg} positions (colours). The algorithm produces for each draw a set of n possible position numbers ranging from 1 to n_{arg} . Taking $n=4$ as an example, the first such set is $[1,1,1,1]$ indicating that the simulated sample consists of 4 values pertaining to position 1 and the last such set is $[n_{\text{arg}}, n_{\text{arg}}, n_{\text{arg}}, n_{\text{arg}}]$, indicating that the simulated sample consists of 4 values pertaining to position n_{arg} . The resulting sample number $M_{\text{total}} = (n_{\text{arg}} + n - 1)! / (n! (n_{\text{arg}} - 1)!)$. The positions are then interpreted as mid values of intervals such that the integral for a chosen PDF over each such interval takes the value $1/n_{\text{arg}}$. The advantage is, that the set of positions needs to be computed only once. The values, *i.e.* the mid values, assigned to the positions need to be calculated for any PDF only once, too. The method works quite well if $n_{\text{arg}} \geq 81$. But the computing times increase beyond those needed when using Monte Carlo. We used this approach successfully to check Monte Carlo results for the PDFs for sample minimum and maximum.

4. Systematic comparison of t -distribution and bootstrap approach

We used Monte Carlo for simulating the samples of repeated measurements and the total bootstrap for mimicking the “PDF” for the sample mean and computed the probabilities for violating the t -criterion and the b -criterion, *i.e.* p_{t-} , p_{t+} , p_{b-} and p_{b+} , for sample sizes ranging from $n=4$ to $n=8$. The results are shown in Tables 4 and 5. By virtue of the symmetry of the problem one expects that $p_{t-}=p_{t+}$ and $p_{b-}=p_{b+}$. This was confirmed by the calculation. The uncertainty associated with the values shown, *e.g.* $u((p_{t-}+p_{t+})/2)$, is at most 0.01. The uncertainties have been determined by means of splitting any Monte Carlo run in 100 sub-runs.

From theory one expects that the t -criterion is only violated in 2,5% of all samples, if the parent PDF is a Gaussian. This was confirmed by the calculations; see Table 4 (upper part). The Table shows also that this is not true for other parent PDFs and that a trend is evident from the parent PDF G to U . The deviation decreases as expected with increasing sample size.

It is known from theory that the b -criterion should only be violated in 2,5% of all samples for any parent PDF, if the sample size is large. The results presented here for small sample sizes show strong deviations and exhibit a trend that is somehow inverse to the one seen with the t -criterion; see Table 4 (bottom part). Some explanation for these observed trends can be seen in Figure 5, which shows the distributions of $t_{0,025}$ and $b_{0,975}$ for the parent PDFs G and U , and from Table 6.

Table 4. Computed probability for violating the t -criterion (upper part) and the b -criterion for various parent-PDFs and sample sizes. The numbers presented are $(p_{t-}+p_{t+})/2$ or $(p_{b-}+p_{b+})/2$ multiplied by 100%. The first column indicates the parent-PDF and the other columns pertain to sample sizes indicated in the first row.

t -criterion	4	5	6	7	8
G	2.51	2.49	2.50	2.51	2.49
T	2.60	2.64	2.65	2.66	2.62
R	3.58	3.28	3.05	2.92	2.80
U	4.38	3.67	3.20	2.93	2.85
b -criterion	4	5	6	7	8
G	9.58	7.67	6.93	6.19	5.62
T	9.64	7.72	6.92	6.14	5.55
R	9.25	7.22	6.38	5.62	5.07
U	8.92	6.75	5.84	5.20	4.73

Figure 5 shows clearly that the distributions for $t_{0,975}$ and $b_{0,975}$ pertaining to the parent PDF G are similar to a Gaussian. The distributions pertaining to the parent PDF U are asymmetric. This asymmetry decreases with increasing sample size.

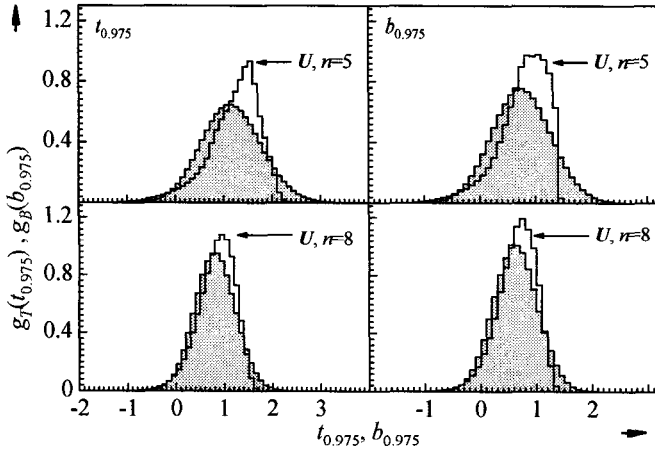


Figure 5. Distribution of $t_{0.975}$ (left figures) and $b_{0.975}$ (right figures) for sample sizes $n=5$ (top) and $n=8$ (bottom) for the parent PDFs G (shaded) and U .

Columns 6 and 7 in table 5 show that the distributions for $t_{0.975}$ have a larger variance than those for $b_{0.975}$. The difference in variance decreases with increasing sample size, too. Columns 2 and 3 in table 5 show that the expanded uncertainty as calculated with the bootstrap is always smaller than the “width” of the sample, i.e. $x_n - x_1$. Columns 4 and 5 in table 5 show that the t -distribution approach uses information, i.e. the parent PDF is G , that *can not be inferred* from the measured samples.

Finally, Figure 6 shows the distribution of the maximal value x_{\max} in a given sample. The distribution is termed S and is shown for the parent PDF G and R . Here R and not U is shown, since the curve for U would dwarf that for G . The properties of R discussed below are in general more pronounced for U . The distribution of the minimal value x_{\min} are obtained by mirroring on the $\xi=0$ axis. The shapes of S using G as parent PDF are not Gaussian but exhibit a long tail to high values. The shapes of S using R as parent PDF are even less Gaussian and exhibit a cut off at $\xi=3^{1/2}$. From this one can infer, that the mean width of the intervals $[x_{\min}, x_{\max}]$ is larger for G than for R and even more so for U . On the other hand, the sample variance is the same for all parent PDFs. This explains the values for $(t_+ - t_-)/(x_n - x_1)$ $(t_+ - t_-)/(x_n - x_1)$ and $(b_+ - b_-)/(x_n - x_1)$ in Table 5.

Table 5. Computed expectation values of ratios of intervals extended by a simulated sample $[x_n-x_1]$, the expanded uncertainties computed using the t -distribution and the bootstrap approach, i.e. $[t_{0.975}-t_{0.025}]$ and $[b_{0.975}-b_{0.025}]$ For more details see the main text.

	$(b_{0.975}-b_{0.025})/(x_n-x_1)$		$(t_{0.975}-t_{0.025})/(x_n-x_1)$		$(t_{0.975}-t_{0.025})/(b_{0.975}-b_{0.025})$	
	<i>G</i>	<i>U</i>	<i>G</i>	<i>U</i>	<i>G</i>	<i>U</i>
4	0.77	0.78	1.43	1.48	1.87	1.88
5	0.65	0.68	1.01	1.07	1.56	1.58
6	0.55	0.59	0.79	0.86	1.45	1.47
7	0.49	0.54	0.66	0.73	1.36	1.37
8	0.44	0.50	0.57	0.65	1.30	1.30

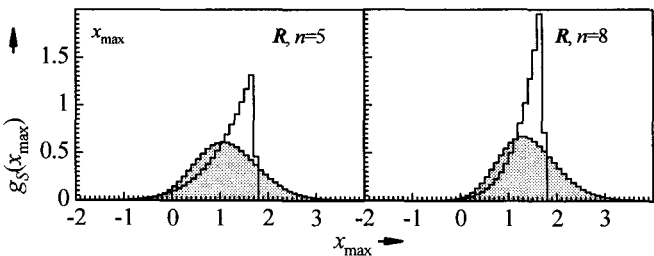


Figure 6. Distribution of x_{\max} , i.e. the largest value of a given sample, using the parent PDFs G (shaded) and R as computed using the total bootstrap on M samples of measured values simulated by Monte Carlo. The sample sizes are $n=5$ (left) and $n=8$ (right). The largest possible value for $\xi \in \mathbf{R} = 3^{1/2}$. In case of U or T one obtains $2^{1/2}$ or $6^{1/2}$, respectively.

5. Summary and Conclusions

In practical metrology, the number of repeated measurements is often small due to the experimental conditions or the high costs of measurements. Generally, sample sizes of $n < 30$ can be considered to be too small to provide a reliable determination of uncertainties.

The performance of the bootstrap and t -distribution approach for the determination of the expanded uncertainty has been compared. A fast algorithm for enumerating the *total* bootstrap and the *total* median and tables for the latter for sample sizes from $n=4$ to $n=17$ have been presented.

The fast algorithm allowed a systematic comparison of the two approaches for the determination of the expanded uncertainty to be performed within reasonable computing times. Monte Carlo was only used to simulate measured data, here termed samples, whereas the bootstrap samples were computed using

tables for the total bootstrap. The sample sizes studied range from $n=4$ to $n=8$, i.e. the degrees of freedom from $\nu=3$ to $\nu=7$. The dependence of the expanded uncertainty calculations on the probability distribution function (PDF) for the measured quantity was also examined by considering four parent PDFs, a Gaussian (*G*), rectangular (*R*), triangular (*T*) and a U-shaped PDF (*U*), that are of relevance in practice.

To conclude, the *t*-distribution approach produces reliable expanded uncertainties if the parent PDF is a Gaussian; however, the information on the parent PDF cannot be inferred from the measured data. For other parent PDFs it produces intervals that contain the mean value with a probability of less than 95 %. The difference depends strongly on the parent PDF and decreases with increasing sample size.

The bootstrap approach produces intervals that contain the mean value with a probability of less than 95 %. The difference decreases with increasing sample size and is larger for *G* and smaller for *U*.

The interval $[t_{0.025}, t_{0.975}]$ is much wider than $[b_{0.025}, b_{0.975}]$, especially for small sample sizes and therefore overly conservative.

In this paper only symmetric parent PDFs have been studied. It is an open question how well the bootstrap works for small sample sizes when the parent PDF is of skewed type that also occurs in metrological applications.

Acknowledgments

We are grateful to Franco Pavese for elucidating discussions.

References

1. Guide to the Expression of Uncertainty in Measurement, International Organization for Standardization, 1995.
2. P. Ciarlini and F. Pavese, *Application of special reduction procedures to metrological data*, Numerical Algorithms **5**, p 479 (1993)
3. I. Lira and W. Wöger, *Bayesian evaluation of the standard uncertainty and coverage probability in a simple measurement model*, Meas. Sci. Technol. **12**, p 1172 (2001).
4. B. Efron and R. J. Tibishirani, *An Introduction to the Bootstrap*, CRC Press LLC, Boca Raton, (Florida) USA (1998).
5. M. Cox and E. Pardo, *The total median and its uncertainty*, In *AMCTM V, Series on Advances in Mathematics for Applied Sciences*, **57**, Ciarlini, Cox, Filipe, Pavese, Richter Eds., World Scientific, Singapore, p 106.

RECURSIVE AND PARALLEL ALGORITHMS FOR APPROXIMATING SURFACE DATA ON A FAMILY OF LINES OR CURVES *

G. ALLASIA

*University of Turin,
Department of Mathematics,
Via Carlo Alberto 10, 10123 Torino, Italy
E-mail: giampietro.allasia@unito.it*

We consider the problem of approximating a continuous real function known on a set of points, which are situated on a family of (straight) lines or curves on a plane domain. The most interesting case occurs when the lines or curves are parallel. More generally, it is admitted that some points (possibly all) are not collocated exactly on the lines or curves but close to them, or that the lines or curves are not parallel in a proper sense but roughly parallel. The scheme we propose approximates the data by means of either an interpolation operator or a near-interpolation operator, both based on radial basis functions. These operators enjoy, in particular, two interesting properties: a subdivision technique and a recurrence relation. First, the recurrence relation is applied on each line or curve, so obtaining a set of approximated curves on the considered surface. This can be done simultaneously on all the lines or curves by means of parallel computation. Second, the obtained approximations of the surface curves are composed together by using the subdivision technique. The procedure gives, in general, satisfactory approximations to continuous surfaces, possibly with steep gradients.

1. Introduction

We consider the problem of approximating a continuous function f from R^2 to R known on a set of points S_n , which are situated on a family of (straight) lines or curves on a domain $D \subset R^2$, convex and bounded. The most interesting case occurs when the lines or curves are parallel. In general, we can admit that some points (possibly all) are not collocated exactly on the lines or curves but close to them, or that the lines or curves are not parallel in a proper sense but roughly parallel. Clearly there is a structure to such data, but it is not required that the data distribution on each line

*This work is supported by University of Torino.

or curve shows a special regularity, that is, the points can be irregularly spaced and in different positions on each line or curve. A frequent feature of this kind of data, often called track data, is that two points which are adjacent to each other along a given track are much closer together than points on different tracks.

As a matter of fact, in several applied problems the function values are known along a number of parallel lines or curves in the plane, as in the case of ocean-depth measurements from a survey ship or meteorological measurements from an aircraft or an orbiting satellite. These data are affected by measurement errors and, generally, taken near to rather than precisely on straight or curved tracks, owing to the effects of disturbing agents (wind, waves, etc.). Moreover, similar geometric structures of data arise in many physical experiments which investigate the dependence of one variable on another for each of a number of discrete values of a third variable.

Several methods (see, e.g., [4, 5, 8, 10, 11] and references therein) have been proposed to solve some restricted formulations of the considered problem by using different approximation techniques (approximation or interpolation, as opposed to approximation) and tools (tensor-product splines, least squares, radial basis functions, Chebyshev polynomials of the first or second kind, etc.).

The present paper concerns mainly the approximation of a function whose coordinates are specified at points on or near families of parallel or roughly parallel lines or curves. A variety of boundaries to the data are admissible, so it is not required to transform boundaries to yield rectangular regions. Functions to be fitted are continuous, possibly with steep gradients. It is noteworthy that our approximation method does not require in itself any data structure, namely, it applies to scattered data. So, it can be applied to data on a family of irregularly distributed polygonal curves such as the bathymetry ship tracks in the Marianas trench considered in [5].

The approximation is obtained by means of a near-interpolation operator or an interpolation operator, both based on radial basis functions. Precisely, the near-interpolant is obtained by introducing a shape parameter in a cardinal radial basis interpolant, so that the resulting formula will not be interpolating in a strict sense, but tends to the interpolation formula as the parameter goes to zero [6, 7]. On the other hand, the shape parameter makes the formula more flexible and allows extension of the set of functions to be approximated. In Section 2 we define jointly the near-interpolation operator and the interpolation operator, giving a constructive way to ob-

tain them. Moreover, we point out two basic properties of these operators, which can be easily used to construct algorithms for iterative and parallel computation. Section 3 is devoted to discuss practical aspects of the application of the operators to approximating surface data on a general family of lines or curves. Section 4 considers special devices to be adopted when the lines or curves are parallel or roughly parallel. Numerical experiments are carried out in Section 5 by considering classical test functions, defined on two sets of track data (lines and exponential curves), and also on two other sets, obtained by perturbing the previous ones. Moreover, a practical application of the operators is showed, which concerns the reconstruction of a surface of a stone ashlar in a Roman theatre. At last, in Section 6 some conclusions are drawn.

2. Interpolation and Near-Interpolation Operators

Let $S_n = \{x_i, i = 1, 2, \dots, n\}$ be a set of distinct points (nodes) in a domain $D \subset R^2$ with associated real values $\{f_i, i = 1, 2, \dots, n\}$, and let $\alpha(x, y; r)$, with $x, y \in D$ and $r \geq 0$, be a continuous positive real function such that

$$\lim_{r \rightarrow 0} \alpha(x, y; r) = \alpha(x, y),$$

where $\alpha(x, y) > 0$ if $x \neq y$, $\alpha(x, y) = 0$ if $x = y$. Define the basis functions

$$g_j(x; r) = \frac{\prod_{k=1, k \neq j}^n \alpha(x, x_k; r)}{\sum_{h=1}^n \prod_{k=1, k \neq h}^n \alpha(x, x_k; r)}$$

and the operator

$$F_{S_n}(x; r) = \sum_{j=1}^n f_j g_j(x; r) = \sum_{j=1}^n f_j \frac{\prod_{k=1, k \neq j}^n \alpha(x, x_k; r)}{\sum_{h=1}^n \prod_{k=1, k \neq h}^n \alpha(x, x_k; r)},$$

or equivalently

$$F_{S_n}(x; r) = \sum_{j=1}^n f_j \frac{1/\alpha(x, x_j; r)}{\sum_{h=1}^n 1/\alpha(x, x_h; r)}, \quad (1)$$

$$F_{S_n}(x_i; 0) = f_i, \quad i = 1, \dots, n.$$

If $r = 0$, then $F_{S_n}(x) \equiv F_{S_n}(x; 0)$ is an **interpolation operator** to the underlying function $f(x)$ at the nodes x_i , ($i = 1, \dots, n$), and the $g_j(x) \equiv g_j(x; 0)$ are cardinal, that is, $g_j(x_i) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. On the other hand, if $r \neq 0$, the operator $F_{S_n}(x; r)$ is no longer

interpolating, but can be considered, for small values of the parameter r , as a **near-interpolation operator**. Note that

$$F_{S_n}(x; r) = F_{S_n}(x) + \sum_{j=1}^n f_j [g_j(x; r) - g_j(x)]. \quad (2)$$

Many choices are possible for the function $\alpha(x, y; r)$. Nevertheless experience suggests to identify α with a radial basis function

$$\alpha(x, y; r) = \phi(\|x - y\|^2 + r),$$

where $\|\cdot\|$ is a convenient norm, generally, the Euclidean norm $\|\cdot\|_2$. Two of the standard functions commonly used in radial basis function approximation are

$$\phi_1 = \|x - y\|_2^2 + r, \quad \phi_2 = \exp[\rho\|x - y\|_2^2 + r], \quad \rho > 0. \quad (3)$$

The operator $F_{S_n}(x; r)$ just defined, including its special case $F_{S_n}(x)$, enjoys many interesting properties [1]; we remark on two of them. A **sub-division technique** can be applied, achieving noteworthy results, very well suited for parallel computation [2]. Let us make a partition of the set of nodes S_n into q subsets S_{n_j} , so that the j th subset, ($j = 1, \dots, q$), consists of the nodes $x_{j1}, x_{j2}, \dots, x_{jn_j}$, with $n_1 + n_2 + \dots + n_q = n$, and the values f_{jk_j} , ($j = 1, \dots, q; k_j = 1, \dots, n_j$), correspond to the nodes x_{jk_j} . Then $F_{S_n}(x; r)$ can be rewritten in the form

$$F_{S_n}(x; r) = \sum_{j=1}^q F_{S_{n_j}}(x; r) \frac{A_j}{\sum_{j=1}^q A_j}, \quad (4)$$

where

$$A_j = \sum_{k_j=1}^{n_j} 1/\alpha(x, x_{jk_j}; r).$$

As a particular case, i.e., for $q = 2$, the following *multistage procedure* works very well. In the first stage, a given set of nodes $S_{n_1} = \{x_i, i = 1, \dots, n_1\}$ is considered and the corresponding operator $F_{S_{n_1}}(x; r)$ is evaluated. In the second stage, it is required to enlarge the considered set S_{n_1} , taking the union of it and another set of nodes $S_{n_2} = \{x_j, j = 1, \dots, n_2\}$. Now the operator on the union set $S_{n_1} \cup S_{n_2}$, with $S_{n_1} \cap S_{n_2} = \emptyset$, can be obtained simply by evaluating $F_{S_{n_2}}(x; r)$, related to the added set S_{n_2} , and using the relation

$$F_{S_{n_1} \cup S_{n_2}}(x; r) = \frac{F_{S_{n_1}}(x; r)A_1 + F_{S_{n_2}}(x; r)A_2}{A_1 + A_2},$$

where

$$A_1 = \sum_{i=1}^{n_1} 1/\alpha(x, x_i; r), \quad A_2 = \sum_{j=1}^{n_2} 1/\alpha(x, x_j; r),$$

and A_1 is known.

Setting $S_{n_1} = S_n$ and $S_{n_2} = \{x_{n+1}\}$, we have that an additional node x_{n+1} can be added to the original set S_n by simply combining an extra term with the original formula, as expressed in the *recurrence relation*

$$F_{S_{n+1}}(x; r) = \frac{F_{S_n}(x; r)A_n + f_{n+1} 1/\alpha(x, x_{n+1}; r)}{A_n + 1/\alpha(x, x_{n+1}; r)}, \quad (5)$$

where

$$A_n = \sum_{k=1}^n 1/\alpha(x, x_k; r).$$

The same formula provides also the tool to eliminate one by one points from the original set of data. It is sufficient to solve (5) with respect to $F_{S_n}(x; r)$. The possibility of adding or eliminating points from the original data set may be useful in some cases (optimal subset selection, interactive graphics, etc.).

3. Basic Procedure

To approximate surface data on a general family of lines or curves using the operator $F_{S_n}(x; r)$ we can proceed in several ways, which depend mainly on data structure, data errors, and expected results. Omitting details pertinent to specific situations, the basic procedure could be sketched as follows.

Step 1. The lines or curves are numbered from 1 to q .

Step 2. The set S_n is partitioned into q subsets S_{n_j} , ($j = 1, 2, \dots, q$), so that the points of S_{n_j} belong to the j th line or curve, namely, are on or near the line or curve.

Step 3. All the points of S_{n_j} , i.e., $x_{j1}, x_{j2}, \dots, x_{jn_j}$, are ordered with respect to a direction on each line or curve.

Step 4. The recurrence relation (5) for $F_{S_{n_j}}(x; r)$ is applied on the ordered set S_{n_j} for any point $x \in D$.

Step 5. Approximations obtained for the j th line or curve, with $j = 1, 2, \dots, q$, are composed together by the subdivision technique (4), so obtaining an approximation to the underlying surface at the point x .

We give some comments on the use of the procedure:

(a) Step 4 can be carried out simultaneously on subsets of the family of lines or curves by means of parallel computation. Suppose we consider $q = kp + r$ lines or curves, where $p + 1$ is the number of processors available and $0 < r \leq k$. Then the first k lines or curves are assigned to the first processor, the second k lines or curves to the second processor, and so on up to the last processor which handles r lines or curves. More information on both serial and parallel algorithms on different architecture can be found in [2]. There it is shown in particular that under condition of well balanced workload the speed-up factor of the parallel computation is approximately equal to the number of processors.

(b) If we let x vary on a suitable discrete set of points on the j th line or curve, ($j = 1, 2, \dots, q$), we obtain an approximation to that curve on the surface whose projection is the j th line or curve. Obtaining good approximations to all curves on the surface corresponding to the plane lines or curves of the given family can be in itself of considerable interest in many applications, as an example to optimize information on topographical contour lines in the construction of a digital elevation model.

(c) Step 5 must be completely executed when the operator is evaluated at the points of a regular grid. Approximating the function values on a grid is requested to represent the surface on a computer display, but it can also be useful in other cases. In fact, remapping the original data into a rectangular mesh by means of $F_{S_n}(x; r)$ can represent the preliminary operation in order to apply techniques based on tensor-product approximation (see, e.g., [11]).

Now we discuss briefly how to restrict two crucial shortcomings in the approximation by $F_{S_n}(x; r)$, namely, the occurrence of flat spots at the nodes and the dependence of the operator on all the nodes. The aim can be reached by using local approximations and mollifying functions (see, e.g., [1]).

To avoid the generally undesirable property that a flat spot occurs at each data point, the use of information about derivatives, either given or generated from data, is recommended and may result in substituting in (1) each functional value f_j with any approximation $L_j(x)$ to $f(x)$ in x_j such that $L_j(x_j) = f_j$, ($j = 1, 2, \dots, n$). In particular, $L_j(x)$ can be the truncated Taylor expansion of $f(x)$ about $x = x_j$, up to derivatives of a certain order evaluated at the point x_j ; however, the technique calls for additional derivative values that are not normally available as data. A more practical solution is to determine a local approximant $L_j(x)$ to $f(x)$ at the point x_j , obtained by means of a moving least squares method using

weights with reduced compact support.

Another computational problem is due to $F_{S_n}(x; r)$ being global, that is, each interpolated value depends on all the data. When the number of data is very large, the calculation of the operator becomes proportionately longer and, eventually, the method will become inefficient or impractical. If the weighting function means that only nearby data sites are significant in computing any interpolated value, a considerable saving in computation could be effected by eliminating calculations with distant data sites. Several authors have suggested localizing schemes in such a way as to obtain a weighting function which is zero outside some disk of suitable radius centered at each node (see [14]). Hence, to localize $F_{S_n}(x; r)$ one can multiply the functions $1/\alpha(x, x_j; r)$ in (1) by the so-called mollifying functions $\tau_j(x, x_j)$, ($j = 1, 2, \dots, n$), which are nonnegative, have local supports in some appropriate sense and satisfy $\tau_j(x_j, x_j) = 1$, as for example the cut functions

$$\tau_j(x, x_j) = (1 - \|x - x_j\|_2^2 / r_j^2)_+, \quad (6)$$

where r_j is the radius of the circle of support at the point x_j .

An interesting alternative to cut functions is offered by the function

$$\tau(t) = \begin{cases} -2^{3n}t^3 + 3 \cdot 2^{2n}t^2 - 3 \cdot 2^n t + 1, & \text{if } 0 \leq t \leq 1/2^n, \\ 0, & \text{if } t > 1/2^n, \end{cases} \quad (7)$$

where $t = \|x\|_2^2$. In fact, we have $\tau(0) = 1$ and $\tau(1/2^n) = 0$; the function is convex and its tangent plane at $t = 1/2^n$ is horizontal; the localizing effect increases with n . Localizing functions like $\tau(t)$, possibly with different orders of continuity, may represent an alternative choice to families of localizing functions based on cut functions (see, e.g., [17]).

Summing up, we are led to consider the following general form of (1)

$$\begin{aligned} \tilde{F}_{S_n}(x; r) &= \sum_{j=1}^n L_j(x) \frac{\tau_j(x, x_j)/\alpha(x, x_j; r)}{\sum_{h=1}^n \tau_h(x, x_h)/\alpha(x, x_h; r)}, \\ \tilde{F}_{S_n}(x_i; 0) &= f_i, \quad i = 1, \dots, n. \end{aligned} \quad (8)$$

4. Special Devices

The considered basic procedure, generally working with an operator in the form (8), can improve in order to handle at best particular data structures, mainly with regard to using mollifying functions and local approximations. A number of cases are worthy to be discussed separately.

Parallel lines. It is not restrictive to suppose that the lines are parallel to the abscissa axis, so that the distance between any pair of lines is immediately determined. Then, evaluating the operator at a point x , only the nodes belonging to a convenient disc centered at x are considered, so restricting the use of the mollifying function with a remarkable reduction of computational effort. This is easily done by calculating directly a relatively small number of distances, in order to find the lines nearest to x and the points on these lines nearest to the projections of x on the lines.

To construct a local approximant in each node of the disc, we distinguish between two different situations. If the nodes lie on the lines and are much closer together than the lines themselves, then a univariate discretized Taylor formula of order two can be used. Precisely, the first and second derivatives at a node are estimated by means of numerical differentiation formulas, involving three consecutive nodes. Otherwise, if the nodes are not exactly lying on the lines, one can consider their projections on the lines themselves, assuming that the nodes would be on the lines in the absence of noise. But it is more cautious to find a second degree polynomial which represents the least squares approximation at each considered node obtained by means of a number of closest nodes.

Parallel curves. As in the previous case, the distance between any pair of curves can be easily determined and only the nodes belonging to a convenient disc centered at x , the point of interpolation or near-interpolation, can be considered, so that also now the use of the mollifying function can be restricted.

To construct a local approximant in each node of the disc, a univariate discretized Taylor formula of order two can still be used, provided that the curves are sufficiently well approximated by a polygonal curve, the nodes lie on the curves, and these ones are not too close together. However, this requires particular care near the vertices of the polygonal curve. An alternative way of more general application consists in finding a least squares polynomial as described above, and this works well also when the nodes are not exactly lying on the curves.

Roughly parallel lines or curves. Once again, a convenient disc centered at x can be determined with relative ease and the use of the mollifying function can be restricted.

To construct a local approximant in each node of the disc, approximation by a least squares polynomial is recommended both for nodes on or near the lines or curves.

Interpolation versus near-interpolation. In many applications, function values are subject to errors; hence it is not appropriate to interpolate the function at the data in the sense of exact matching, but it seems more appropriate to approximate the function or, more precisely, to consider near-interpolation. Data requiring near-interpolation, ranging from points on a rectangular mesh to arbitrarily scattered, occur in virtually every field of science and engineering. Sources include both experimental results (experiments in chemistry, physics, engineering) and measured values of physical quantities (meteorology, oceanography, optics, geodetics, mining, geology, geography, cartography), as well as computational values (e.g., output from finite element solutions of partial differential equations). In particular, since the problem of constructing a smooth surface from a given set of lines or curves appears in many instances in geophysics and geology, the near-interpolation operator can meet this requirement.

The parameter r in the near-interpolation operator $\tilde{F}(x; r)$ in (8) has the effect that, in general, the gradient of the rendered surface is not zero at the nodes. As a consequence, the surface is considerably smoother than for $r = 0$. However, if r is too small, the first derivatives of $\tilde{F}(x; r)$ are highly oscillating and their values are nearly zero. Clearly, the goal is to choose an "optimal" value of r , such that $\tilde{F}(x; r)$ does not present flat spots, but, in the same time, it maintains a sufficient computational accuracy, in particular at the nodes (see [3] for a discussion on the choice of parameter values).

It should be noted that near-interpolation errors $F_{S_n}(x_i; r) - F_{S_n}(x_i)$ at the nodes in (2) must not be confused with the unknown errors which affect the corresponding function values f_i . However, it is reasonable to arrange matters so that near-interpolation errors and errors on f_i are quantities of the same order.

On the other hand, in some cases, it may be interesting to consider interpolation as opposed to approximation. As an example, in [4] the construction of tensor-product spline interpolants to various types of data structure is discussed. Such interpolants can be constructed efficiently and directly for data on a rectangular mesh, but also for data on a family of parallel lines and data lying close to a family of parallel lines.

5. Some Tests and a Practical Application

Carrying out numerical experiments on a number of test functions, we considered two sets of data, actually given in different domains but mapped

for uniformity into the square $[0, 1] \times [0, 1]$. The former set arises from measurements by photogrammetric survey on a deteriorated face of an ancient ashlar (see below). The latter is obtained by considering the family of curves [16]

$$y_p(x) = 0.5(p+1) \exp[0.1(x+1)] + 0.05(x+1) \exp[0.5(p+1)-1] + 0.5(p-1),$$

where $p = -1 + kh$, ($k = 0, \dots, m_2 - 1$), and $h = 2/(m_2 - 1)$. On each curve y_p a set of equispaced abscissas $x_j^{(p)}$, ($i = 1, \dots, m_1^{(p)}$), is considered. The values of $m_1^{(p)}$ and m_2 are chosen suitably. Moreover, from the sets just defined, we obtained two other sets by slightly perturbing both the coordinates of the data points and the distances between pair of contiguous lines or curves.

Then, we generated the function values at the points of the considered sets for each of Franke's six test functions [12, 13], using 703 data points for the photogrammetric lines and 635 for the exponential curves in the case of the original data sets, and almost the same numbers in the case of the perturbed data sets. The basic procedure and the special devices, described in Sections 3 and 4 respectively, have been applied to reconstructing the test surfaces by the operator $\tilde{F}_{S_n}(x; r)$ in (8), considering all the possible options, namely, $r \neq 0$ vs. $r = 0$, ϕ_1 vs. ϕ_2 in (3), and $\tau_j(x, x_j)$ vs. $\tau(t)$ in (6) and (7) respectively. In every case the achieved performances agree as regards to accuracy with those described in literature or expected from analogy, but the procedure allows an impressive improvement in flexibility and computer time.

The near-interpolation operator can be considered as a tool to reconstruct a lateral surface of a stone ashlar in the Roman theatre in Aosta (Aosta Valley, Italy). Over the centuries, the ashlar has been seriously damaged from atmospheric agents, suffering a loss of material so that the considered surface is actually very rough. The surface reconstruction represents a step in implementing the software for a dynamical tridimensional representation of parts of the ancient building in order to plan restoration strategies. The problem has been considered in details in [9] by using a cross-sectional technique and nonuniform rational B-splines (see also [15] and [3], where a different technique is employed).

On a rectangular region of a lateral face of the ashlar, the coordinates (x, y, z) of each mesh point of a grid have been measured by photogrammetric survey, where (x, y) gives the position of the point in the original plane of the face, and z is a measure of the variation with respect to this plane, that is, z represents a measure of damage. The dimensions in millimetres of

the region are $60 \leq x \leq 960$ and $81.6 \leq y \leq 531.6$, whereas $-160 \leq z \leq 0$. The points lie on 19 lines and each line contains 37 points.

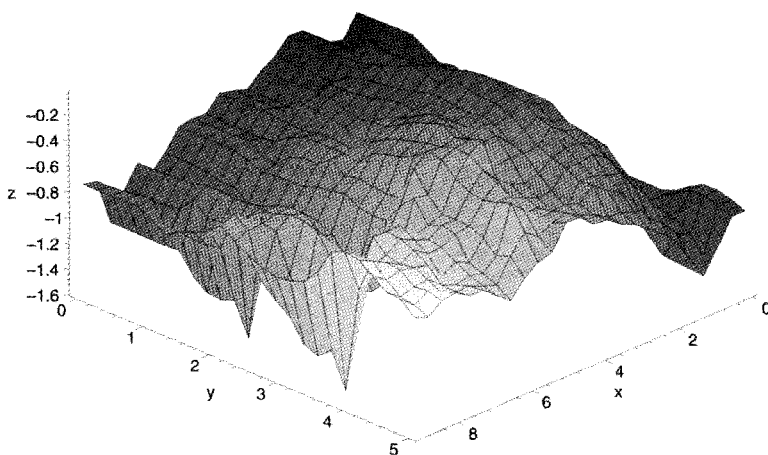


Figure 1. Near-interpolation operator reconstruction of the ashlar surface.

From the data many surface reconstructions have been obtained by $\tilde{F}_{S_n}(x; r)$, considering the several options illustrated above. To shorten, it results convenient to use the near-interpolation operator, to set simply $M_j(x) = f_j$, because the surface is not smooth, and to adopt the tentative parameters values suggested by experience on test functions. In practice, the surface in Figure 1 is a convenient representation of the actual surface on the ashlar.

6. Conclusion

Given a family of lines or curves on a plane domain, possibly parallel or roughly parallel, we consider the problem of approximating a continuous real function, known on a set of points situated on or near the lines or curves. The procedure we propose approximates the data by means of either an interpolation operator or a near-interpolation operator, both based on radial basis functions. Since the operators enjoy some noteworthy properties, the procedure succeeds in reconstructing surfaces with a satisfactory approximation, but especially shows a considerable flexibility and an impressive reduction in computer time, thus offering a very convenient instrument of application.

Acknowledgments

I would like to thank the referees for their useful comments.

References

1. G. Allasia, in *Approximation Theory, Wavelets and Applications*, ed. S. P. Singh, (Kluwer, Dordrecht, 1995), pp. 1–36.
2. G. Allasia and P. Giolito, in *Surface fitting and multiresolution methods*, eds. A. Le Mèhautè *et al.*, (Vanderbilt Univ. Press, Nashville, TN, 1997), pp. 1–8.
3. G. Allasia and R. Besenghi, Research Report n. 56/2002 Department of Mathematics, University of Turin, submitted.
4. I. J. Anderson, M. G. Cox and J. C. Mason, *Numer. Algorithms* **5** (1993), 193–204.
5. D. Apprato, C. Gout and D. Komatitsch, *Math. Geol.* **34** (7) (2002), 831–843.
6. R. Besenghi and G. Allasia, in *Curves and surfaces fitting: Saint-Malo 1999*, eds. A. Cohen *et al.*, (Vanderbilt Univ. Press., Nashville TN, 2000), pp. 75–84.
7. R. Besenghi, M. Costanzo and A. De Rossi, *Int. J. Comput. Numer. Anal. Appl.* **3** (4) (2003), 419–438.
8. R. E. Carlson and T. A. Foley, *Computers Math. Applic.* **12** (1992), 27–34.
9. P. Ciarlini and S. Cipriani, Quaderno n. 31/1996 IAC–CNR, Roma, Italy, 1–40.
10. A. Crampton and J. C. Mason, in *Advanced mathematical and computational tools in metrology*, Vol. 5, eds. P. Ciarlini *et al.*, (World Scientific, Singapore, 2000), pp. 119–126.
11. T. A. Foley and G. M. Nielson, in *Approximation theory III*, ed. E. W. Cheney, (Academic Press, New York, 1980), pp. 419–424.
12. R. Franke, Report TR-NPS-53-79-003 (1979) Naval Postgraduate School, Monterey, California.
13. R. Franke, *Math. Comp.* **38** (157) (1982), 181–200.
14. R. Franke and G. Nielson, *Intern. J. Numer. Methods Eng.* **15** (1980), 1691–1704.
15. P. Grattoni and M. Spertino, Report 2002 IRITI-CNR, Torino, Italy.
16. J. C. Mason and R. P. Bennel, in A. Bowyer (ed.), *Computer-aided surface geometry and design (Bath, 1990)*, Inst. Math. Appl. Conf. Ser. New Ser. **48**, Oxford Univ. Press, New York, 1994, 33–50.
17. R. Schaback, in *Mathematical methods for curves and surfaces*, eds. M. Dæhlen *et al.*, (Vanderbilt Univ. Press, 1995), pp. 477–496.

PROCESS MEASUREMENT IMPACT ON THE VERIFICATION UNCERTAINTY

J. BACHMANN, J. M. LINARES, S. ARANDA AND J. M. SPRAUEL

*Laboratoire EA(MS)², Université de la Méditerranée, I.U.T., Avenue Gaston berger - F
13625 Aix en Provence cedex 1, E-mail: Linares@iut.univ-aix.fr*

Coordinate measuring machines are now widely used to qualify industrial pieces. Nevertheless, the actual CMM software's usually restrict to the determination of mean values. This is the case for both the characterization of individual surfaces and for the determination of geometrical errors. However, in accordance with quality standards, the uncertainty of each measurement should also be defined. At last CIRP seminar, a new non linear least squares method has been proposed for that purpose to define the error bars of the parameters estimated for each measured surface. These values are deduced from the gap between the measured coordinates and the associated optimized surface. Our new presentation now extends to the propagation of such uncertainties to the determination of ISO1101 tolerances (dimensions and geometrical errors). To illustrate this approach, a specification was inspected on a true real industrial piece, with respect to the ISO 1101 standard. For this industrial application, different measurement procedures were proposed and carried out. The uncertainties of the estimated geometrical errors were then evaluated, showing the influence of the experimental method onto the reliability of the measurement. This example thus demonstrates the need to optimize the inspection process. To conclude our presentation, different aspects of the inspection are finally discussed to improve the verification of ISO 1101 specifications.

1. Introduction

Different research programs were started during a recent CIRP seminar [1] concerning the control of uncertainties in the whole industrialization process of a mechanical unit. A generalized principle of uncertainty was thus presented [2]. In this topic the uncertainty of measurement is also considered. The concept of uncertainty is well known in metrology. It is represented by a statistical parameter associated to the result of a given measurement and is related to the fact that each measurement is altered by random errors, which have to be quantified and reported in all the results or operations which derive from the measurement process.

Generally, only the uncertainties of the parameters that characterize each elementary surface are estimated. A lot of research work has already been done on this topic. Among these studies, some relate to the uncertainties induced by the measuring instrument. Others deal with the data-treatment procedures performed after the measurement process.

In the same context, our paper will show some results obtained through a new data-processing method. A vectorial approach of surfaces was used for that

purpose. Taking account of the significant increase in the performance of computers, no linearization of the equations was however employed, contrary to most of the current algorithms. The results are thus expressed in a global 3D coordinate system, which remains the same for all the analyzed surfaces. The interest of such approach also results in the fact that it allows an automatic calculation of uncertainties. These uncertainties are simply derived from the set of acquired coordinates, which is considered as a statistical sampling of the analyzed surface. The covariance matrix of the parameters that characterize each feature is also estimated, thus allowing defining a confidence interval that bounds the true real surface.

If various methods are proposed to define uncertainties on an elementary surface, the error bars on the evaluation of geometrical deviations are seldom indicated. In accordance with the directives of the ISO TC 213, a new method of propagation of uncertainties is therefore proposed. The interpretation of functional specifications, expressed through ISO 1101 standards, requires calculating distances between different geometrical elements:

- Datum surfaces and datum systems (ISO 5459), which are mathematical surfaces described by the vectorial geometry: point, line or plane.
- Specified surfaces which are generally known only through a set of digitized points.

The calculation of the uncertainties of a given specification comes thus to the evaluation of the error bars of these calculated distances. This operation consists in propagating the covariance matrix of the parameters that characterize each measured elementary surface down to each derived distance [3]. Such propagation is a mechanism that requires data transfer and thus depends on a great number of parameters. The aim of our paper is then to show the influence of the experimental methods onto the final result. Different measurement procedures will therefore be applied to the same work piece. For each experimental configuration, the resulting uncertainties will be derived directly from the set of acquired points.

2. Experimental study

The aim of this paragraph is to define the uncertainties of geometric deviations evaluated to check ISO 1101 type specifications. The tested piece is the part described in Figure 1. Our industrial partner according to his usual experimental procedures digitalized the part. The files of the acquired coordinates were then

post-treated through our algorithms to define the error bars of the measurements.

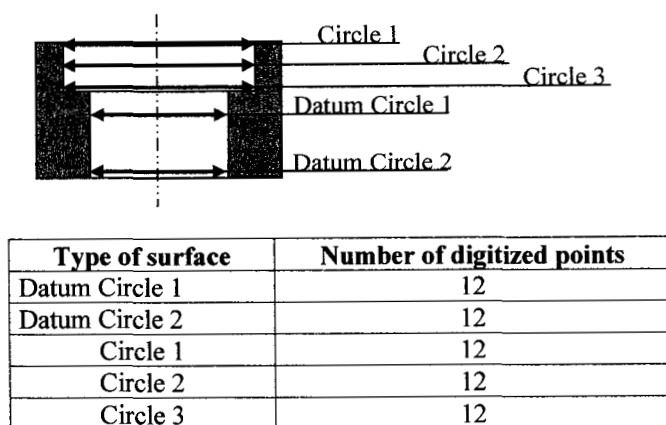


Figure 1. Definition of the studied part and location of the surfaces.

On the tested piece, the coaxially between two cylinders had to be controlled (Figure 2).

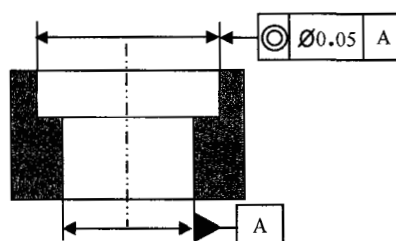


Figure 2. Definition of the specification to be controlled.

For this type of specification, two types of construction can be considered:

- The first solution consists in constructing the line defined by the centers of the two Datum Circles and then to calculate the distances between the centers of the circles that define the specified surface and this built line. The distances obtained in this manner have then to be compared to half of the tolerance imposed by the designer.
- The second solution consists in concatenating the two sets of points defining the Datum circles 1 and 2, to compute the cylinder that fits the resulting data and finally to calculate the distance between the axis of this

cylinder and the centre of the circles which define the axis of the specified cylinder.

Following Figure 3 summarizes both constructions.

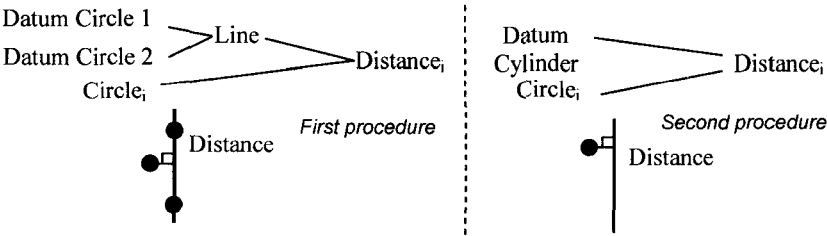


Figure 3. Checking process

These two procedures have been implemented. The results derived from the two constructions are reported in table 1. Figures 4 and 5 show the screen captures of the software developed to perform these operations.

Table 1. Results obtained by the two types of constructions.

Procedure	Circle number	Mean value of distance d_i	Uncertainties
First	1	0,0135	8,1E-03
	2	0,0148	7,2E-03
	3	0,0132	6,7E-03
Second	1	0,0135	4,9E-03
	2	0,0148	4,6E-03
	3	0,0132	4,8E-03

The standard deviations are deduced from the uncertainties of the elementary surfaces. These uncertainties are derived directly from the set of acquired points and are then propagated to the calculated distances. The calculated distances are finally to be compared to half the tolerance imposed by the designer.

It has to be noticed that:

- The mean values of the distances calculated by the two methods are identical,
- The Uncertainty of the distance derived from the first method is almost two times larger than the value obtained through the second procedure.

This demonstrates that the choice of procedure has a quantitative influence on the quality of the result.

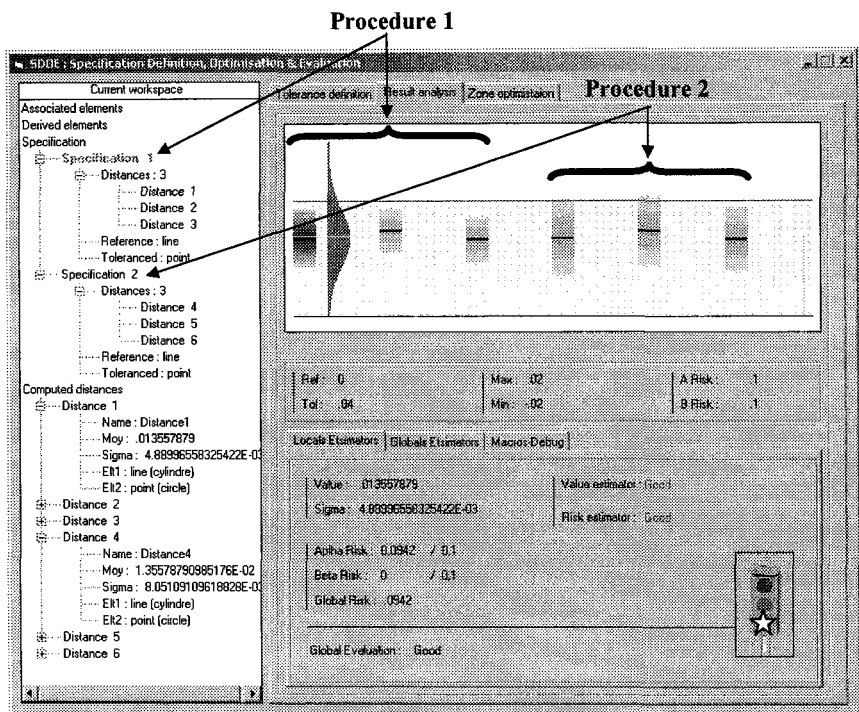


Figure 4. Chart of the results (first procedure)

In that case, indeed, the controlled part will satisfy the tolerance zone of the specification whatever the selected method. On the other hand, with regard to the quality of the results, the uncertainties are two times greater with the first construction than with the second one, while the acquired coordinates are exactly the same for both procedures.

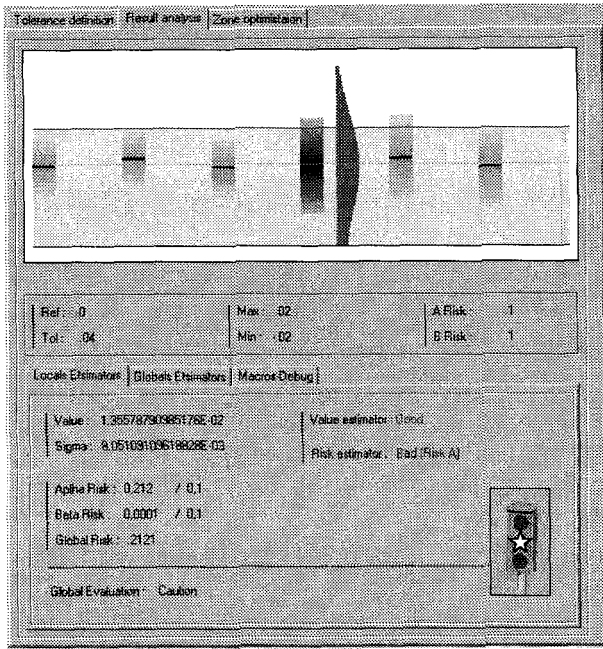


Figure 5. Chart of the results (second procedure)

3. Statistical Confidence Boundary (SCB)

The knowledge of the covariance matrix of the parameters that characterize each feature allows defining a confidence interval that bounds the true real surface. Such interval, called Statistical Confidence Boundary (SCB), may be represented by a three dimensional plot which gives a better understanding of the concept of uncertainty.

3.1. SCB definition

The SCB defines the position and thickness of the interval that delimits the zone where the real surface is included [4] [5]. To illustrate this topic, the confidence interval of a plane is presented, as example. The calculation of this SCB uses the following procedure: a set of points belonging to the ideal plane fitted to the acquired coordinates is considered first. These points are built without any uncertainty. The mean distance between each point and the plane are null by construction, but its variance defines the variability of the feature adjusted to the experimental data. It allows thus calculating, for a fixed confidence level, the

statistical interval that bounds the real surface. This interval defines the SCB [6] [7].

It has to be pointed out that the uncertainty of each calculated distance, and therefore the corresponding thickness of the SCB, are function of the spacing between the center of the plane, where the covariance matrix is known, and the point where the estimation is computed.

According to the Guide for the expression of uncertainties, the following type of propagation law has been used to calculate the variance of each distance:

$$u_c^2(y) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial a_i} \cdot \frac{\partial f}{\partial a_j} \cdot u(a_i, a_j) = \sum_{i=1}^N \left[\frac{\partial f}{\partial a_i} \right]^2 \cdot u^2(a_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial a_i} \cdot \frac{\partial f}{\partial a_j} \cdot u(a_i, a_j) \quad (1)$$

Since the parameters a_i estimated for the feature fitted to the experimental data can be represented by random vectors, the previous expression can be rewritten as follows:

$$\Omega_p(y) = J_y \cdot \Omega(A) \cdot J_y^T \quad (2)$$

Where: $\Omega(A)$ represents the covariance matrix ($n \times m$) of the entries A_i and J_y is the Jacobian matrix defined there by:

$$\begin{cases} \forall i \in \{1 \dots n\}, \forall p \in \{1 \dots n\} \\ \forall a_i^p \in \mathbb{R} \end{cases} \quad (3)$$

$$J_y = \left\{ \frac{\partial y}{\partial a_1^1}, \frac{\partial y}{\partial a_2^1}, \dots, \frac{\partial y}{\partial a_1^p}, \frac{\partial y}{\partial a_2^p}, \frac{\partial y}{\partial a_1^2}, \dots, \frac{\partial y}{\partial a_p^2}, \frac{\partial y}{\partial a_1^n}, \frac{\partial y}{\partial a_2^n}, \dots, \frac{\partial y}{\partial a_p^n} \right\} \quad (4)$$

For a plane normal to direction (Z), and a point, without uncertainty, belonging to this plane, the equation of a quadric is found. As the coordinates of the centre point of the plane and the components of its normal vector are not correlated, this equation is:

$$\text{Var}(d) = X_i^2 \cdot \text{Var}(n_x) + Y_i^2 \cdot \text{Var}(n_y) + K \cdot \text{Var}(O_z) \quad (5)$$

where (X_i, Y_i, Z_i) are the coordinates of the point M_i of the plane, (O_x, O_y, O_z) are the coordinates of the center point of the plane, (n_x, n_y, n_z) are the cosines of its normal vector, K is a constant.

It corresponds to the equation of an elliptic paraboloid. From this expression, the bounding Surfaces of the SCB can be calculated for a fixed confidence ratio k . Figure 7 shows such SCB obtained for a confidence ratio of $k = 2$. The digitized points and the centre of the analyzed surface have been plotted too.

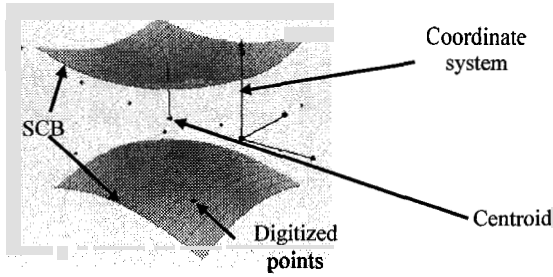


Figure 7: Plane SCB

The SCB of other features like points, lines, etc. can be determined in the same way. In the case of a point, it corresponds to an ellipsoid. The next part of the paper will show that the behavior result can be explained by using the SCB concept.

3.2. Explanation of the results using the SCB concept

The deviations of the two inspection procedures considered in figure 3 can be well defined now, using the SCB concept. The intention of the operator is to control the distance between the center point (M_i) of each circles acquired on the specified surface and the axis of the Datum (O, V). The result of the best fit of the acquired surfaces is represented in figure 8. On the left part of the figure, the SCB's of the datum circles and one section of the specified surface are ellipsoids. On the right part, since the datum is described by a cylinder, the SCB of this surface corresponds to a complex quadric.

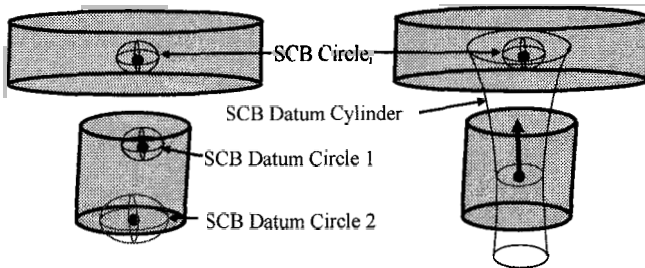


Figure 8: Surfaces SCB

Each distance to control is characterized by its mean value d_i and its uncertainty $u_{c(di)}$, defined by following equations Eq.6 and Eq.7.

$$d_i = \|OM_i \wedge V\| \quad (6)$$

$$u_{c(d_i)} = \sqrt{u^2(M_i) + u^2_{(datum)}} \quad (7)$$

where: d_i is the mean value of the calculated distance, $u_{c d_i}$ is the composed uncertainty of the specification, $u_{(M_i)}$ is the propagated uncertainty of point M_i , $u_{(datum)}$ is the propagated uncertainty of the datum

The figure 9 shows how the composed uncertainty can be deduced directly from the SCB's defined on the Datum and each circle acquired on the specified surface. In the first case the datum is obtained by creating the line defined by the centers of the two Datum circles. The SCB of the resulting item is still a complex quadric. In the second case the datum is defined directly by the cylinder fitted to the acquired points. It is thus clearly demonstrated that the cross section of the datum SCB obtained by the first procedure is larger than the second one. In consequence, the composed uncertainty deduce from the first construction will be much greater than the one calculated through the second procedure.

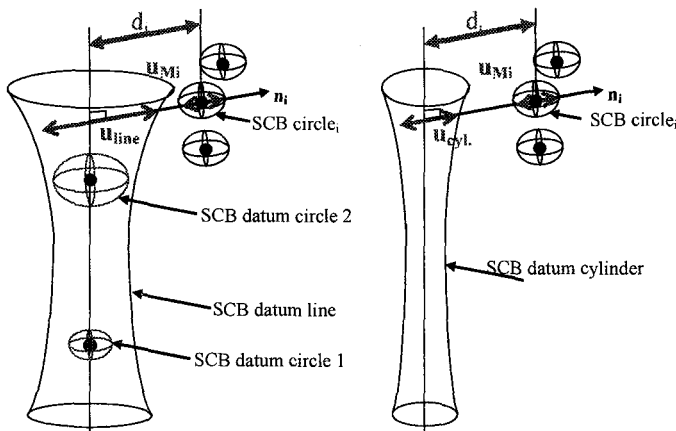


Figure 9: Determination of the composed uncertainty.

4. Conclusion

In this paper, quantitative results highlighted that the procedures selected by the operators of CMM during the design of the measurement process, have a direct consequence on the results of acceptance of the tolerances. The only information usually used to make decision is a mean value. It will never permit

to select between different equivalent procedures the one that gives the most reliable result. For that reason, a generic method has been developed to calculate the uncertainties of each measured specification directly from the set of acquired coordinates. This additional information allows quantifying the risk of acceptance or rejection of controlled parts. In case of indecision, the operator may then reduce the measurement uncertainties either by selecting constructions which will lower the propagation of uncertainties or quite simply by increasing the number of points of the digitized surfaces.

Acknowledgments

The authors would like to thank Professor Pierre Bourdet for his beneficial discussion and advice

References

1. J. Dovmark, "New interesting concepts from ISO/TC 213"; In: *ACMC Annual General Meeting in Ottawa*; Canada 2001.
2. V.Srinivasan, "An integrated view of Geometrical Product Specification and Verification"; In: *Proceedings of the 7th International CIRP Seminar on Computer-Aided Tolerancing*, pp. 7-17; Cachan 2001.
3. J.M. Sprauel, J.M. Linares, P. Bourdet, "Contribution of nonlinear optimization to the determination of measurement uncertainties"; In: *Proceedings of the 7th International CIRP Seminar on Computer-Aided Tolerancing*, pp. 285-291; Cachan 2001.
4. M. Hernla, "Calculation of measuring uncertainty with CMMs under Industrial Conditions"; In: *Proceedings of 3rd CIRP International Seminar on Computer-Aided Tolerancing*; pp. 171-177; Cachan 1993.
5. K. Takamasu, R. Funutani, S. Ozono, "Basic concept of feature-based metrology"; In: *Measurement Vol26*, pp.151-156, 1999.
6. J.M. Linares, J. Bachmann, J.M. Sprauel. and P. Bourdet, Propagation of specification uncertainties in tolerancing, In *Proceedings of the 8th CIRP Seminar on Computer Aided Tolerancing*, Charlotte, pp.301-310 , 28-29 Avril 2003.
7. J.Bachmann, J.M. Linares, J.M. Sprauel and P. Bourdet, Propagation of measurement uncertainties in the inspection of tolerances: Industrial case, In *Proceedings of the 8th CIRP Seminar on Computer Aided Tolerancing*, Charlotte, pp.311-320, 28-29 Avril 2003.

ON THE IN-USE UNCERTAINTY OF AN INSTRUMENT*

WALTER BICH AND FRANCESCA PENNECCHI

Istituto di Metrologia “Gustavo Colonnetti” (IMGC)

Mass Section, Strada delle Cacce, 73

10135 Torino, Italy

E-mail: w.bich@imgc.cnr.it

The in-use uncertainty of an instrument is investigated according to the possible measurement procedures, namely: as a comparator or as a standard. These two alternatives are embedded in a unique model and the uncertainty components due to instrument calibration and noise are discussed. Special attention is given to the case of calibration by fitting with a straight line.

1. Introduction

It is a common opinion that using an instrument as a standard or as a comparator are two fully different experimental techniques for determining the value of a given quantity X . With the first technique, widely adopted in lower-accuracy applications, one uses the calibration curve of the instrument to directly obtain the measurand estimate \hat{X} from a reading Y . The second technique, preferred in highest-accuracy applications, implies comparing the values \hat{X}_1 and \hat{X}_{ref} obtained, through the calibration curve, from two readings Y_1 and Y_2 . These correspond to the measurand X and a reference standard, having a known value X_{ref} as close as possible to the measurand value, being alternatively connected to the instrument. The measurand value is thus estimated by suitably combining the known value of the reference standard with the difference (or ratio) of the estimates given by the comparator. For example,

$$\hat{X}_{\text{comp}} = X_{\text{ref}} + \hat{X}_1 - \hat{X}_{\text{ref}}, \quad (1)$$

where the subscript “comp” stands for “comparison”.

In practice, when using the instrument as a standard, it is good practice to take into account the possible lack of stability of the calibration curve

*Work partially funded under EU SofTools_MetroNet Contract N. G6RT-CT-2001-05061

by means of a “zero” reading, that is, a reading of the instrument response Y_0 , typically different from zero, when a measurand of zero value is fed to its input. To this reading an estimate \hat{X}_0 corresponds, which is subtracted from \hat{X} to form the estimate \hat{X}_d , where the subscript “d” stands for “direct”. Therefore, also in this procedure the measurand estimate involves a difference, that is,

$$\hat{X}_d = \hat{X} - \hat{X}_0, \quad (2)$$

so that the two seemingly different methods can be modelled in similar ways.

The uncertainty provided by method (2) is typically larger than (or at the best comparable with) that obtained with method (1). This means that the difference term $\hat{X}_1 - \hat{X}_{\text{ref}}$ in Eq. (1) has a smaller uncertainty than the corresponding term $\hat{X} - \hat{X}_0$ of Eq. (2). In both cases, the estimate uncertainty can be imagined as the combination of two components, one accounting for repeatability and the second for traceability to SI units, through the reference standard with model (1) and through instrument calibration with model (2). Indeed, as it is well known, in the former difference the dominant uncertainty contribution is random noise of the comparator (and possibly its finite resolution [1, 2]), whereas in the latter both contributions are embedded. Yet, the only distinction between the two difference terms in Eqs. (1) and (2) is quantitative. This suggests that the traceability contribution to the difference uncertainty is somehow proportional to the difference value. We have proved this for a linear calibration curve.

2. The calibration of a measuring instrument

The mathematics underlying the calibration of an instrument is well established. For a complete overview, see, e.g., [3]. Usually, an instrument is calibrated by providing to its input m reference values, or *stimula*, X_i of a given quantity X and by observing the corresponding m *responses* Y_i . The points thus obtained are fitted by a function

$$Y = g(\alpha_0, \alpha_1, \dots, \alpha_{n-1}, X), \quad (3)$$

the parameters α_h of which are to be determined. This estimation problem concerns a vector measurand, that is, the parameter vector α . For example, in the common case of a polynomial function

$$Y = \sum_{h=0}^{n-1} \alpha_h X^h \quad (n \leq m), \quad (4)$$

one can write

$$\begin{aligned} Y_1 &= \alpha_0 X_1^0 + \alpha_1 X_1^1 + \dots + \alpha_{n-1} X_1^{n-1} + \varepsilon_1 \\ Y_2 &= \alpha_0 X_2^0 + \alpha_1 X_2^1 + \dots + \alpha_{n-1} X_2^{n-1} + \varepsilon_2 \\ &\dots \\ Y_m &= \alpha_0 X_m^0 + \alpha_1 X_m^1 + \dots + \alpha_{n-1} X_m^{n-1} + \varepsilon_m \end{aligned} \quad (5)$$

or, in matrix terms,

$$\mathbf{Y} = \mathbf{X}\alpha + \varepsilon, \quad (6)$$

by defining $\mathbf{Y}_{1 \times m}^T = [Y_1 \ Y_2 \ \dots \ Y_m]$, $\alpha_{1 \times n}^T = [\alpha_0 \ \alpha_1 \ \dots \ \alpha_{n-1}]$, and $\mathbf{X}_{m \times n} = [X^0 \ X^1 \ X^2 \ \dots \ X^{n-1}]$. The error vector ε is such that

$$E(\varepsilon) = 0 \text{ and } E(\varepsilon \varepsilon^T) = \mathbf{V}(\mathbf{Y}) = \sigma^2 \mathbf{I}, \quad (7)$$

where $\mathbf{V}(\mathbf{Y})$ is the covariance matrix of \mathbf{Y} and \mathbf{I} is the identity matrix. That means that the responses Y_i are independent and have the same variance σ^2 .

By further assuming that the \mathbf{X} vector has negligible uncertainty compared to that of \mathbf{Y} , the estimated coefficient vector $\hat{\alpha}$ is thus given by a standard least squares adjustment

$$\hat{\alpha} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{Y}. \quad (8)$$

The covariance matrix $\mathbf{V}(\hat{\alpha})$ is obtained from

$$\mathbf{V}(\hat{\alpha}) = \mathbf{J}_Y \mathbf{V}(\mathbf{Y}) \mathbf{J}_Y^T \quad (9)$$

where the *Jacobian* \mathbf{J}_Y is a rectangular $n \times m$ matrix whose generic element \mathbf{J}_{Yij} is $\mathbf{J}_{Yij} = \frac{\partial \alpha_i}{\partial Y_j}$ (see, e.g., [4]).

Application of Eq. (9) to Eq. (8) yields

$$\mathbf{V}(\hat{\alpha}) = u^2(Y_i) [\mathbf{X}^T \mathbf{X}]^{-1}. \quad (10)$$

The analytical form of the estimate $\hat{\alpha}$ and of its covariance $\mathbf{V}(\hat{\alpha})$ depends upon the degree of the polynomial. In the linear case, in which

$$\alpha^T = [\alpha_0 \ \alpha_1] \text{ and } \mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_m \end{bmatrix}, \text{ the well-known solution is [5]}$$

$$\hat{\alpha} = \frac{1}{m \sum (X_i - \bar{X})^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & m \end{bmatrix} \cdot \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}, \quad (11)$$

or

$$\hat{\alpha}_0 = \bar{Y} - \hat{\alpha}_1 \bar{X} \quad (12)$$

and

$$\hat{\alpha}_1 = \frac{\sum (X_i - \bar{X}) (Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}, \quad (13)$$

where $\bar{X} = (\sum X_i) / m$ and $\bar{Y} = (\sum Y_i) / m$. The estimate covariance is

$$\mathbf{V}(\hat{\alpha}) = \frac{u^2(Y_i)}{m \sum (X_i - \bar{X})^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & m \end{bmatrix}. \quad (14)$$

Correlation between slope and intercept can be avoided by a suitable change of variables [1]. In general, this is not possible for higher-degree fittings.

The analytical form of the parameter vector estimate $\hat{\alpha}$ depends on the particular fitting function $Y = g(\alpha, X)$ adopted. This might be as well nonlinear in the parameters, which would imply a nonlinear adjustment.

In several cases the uncertainty of the \mathbf{X} vector is not negligible. The analysis is then more complicated [6, 7], but in any case this further uncertainty contribution is embedded into the uncertainty matrix $\mathbf{V}(\hat{\alpha})$.

3. Using the calibrated instrument

After calibration, the instrument is routinely used in the laboratory. This means that new responses Y_i are observed, from which the corresponding unknown stimula X_i are to be estimated with their uncertainties. The model equation in this application is

$$\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i), \quad (15)$$

where g^{-1} is the inverse function of g in eq. (3) and the vector variable α is now replaced by its estimate $\hat{\alpha}$. From a mathematical viewpoint, the existence of the inverse function is subject to conditions. However, in the range spanned by the calibration stimula, the function g has to be bijective for a meaningful calibration and therefore the inverse g^{-1} exists. The \hat{X}_i values are here arranged in a $\hat{\mathbf{X}}_{q \times 1}$ vector ($q \geq 2$), related to the corresponding \mathbf{Y} vector ($\mathbf{Y}^T = [Y_1 \ Y_2 \ \cdots \ Y_q]$) through the estimated

parameter vector $\hat{\alpha}$. To further simplify the problem, let us arrange the two vectors $\hat{\alpha}_{n \times 1}$ and $\mathbf{Y}_{q \times 1}$ in a column (block) vector $\mathbf{W}_{(n+q) \times 1}$, such that $\mathbf{W}^T = [\hat{\alpha}^T \ \mathbf{Y}^T]$. In order to evaluate the covariance matrix of $\hat{\mathbf{X}}$, we use Eq. (9), here taking the form

$$\mathbf{V}(\hat{\mathbf{X}})_{q \times q} = \mathbf{J}_W \mathbf{V}(\mathbf{W}) \mathbf{J}_W^T. \quad (16)$$

It is worth noting that the propagation law (16) is exact for linear functions only. For a generic function, such as that of Eq. (15), it represents a first order approximation. The covariance matrix $\mathbf{V}(\mathbf{W})_{(n+q) \times (n+q)}$ of \mathbf{W} is an appropriate block matrix constructed with the covariance matrices $\mathbf{V}(\hat{\alpha})_{n \times n}$ and $\mathbf{V}(\mathbf{Y})_{q \times q}$, namely

$$\mathbf{V}(\mathbf{W}) = \begin{bmatrix} \mathbf{V}(\hat{\alpha}) & \mathbf{0} \\ \mathbf{0} & \mathbf{V}(\mathbf{Y}) \end{bmatrix}. \quad (17)$$

Since the vectors $\hat{\alpha}$ and \mathbf{Y} are uncorrelated, the off-diagonal blocks are equal to zero. The lower block, $\mathbf{V}(\mathbf{Y})$, is diagonal, the upper, $\mathbf{V}(\hat{\alpha})$, is not. By further using block matrices, Eq. (16) may be re-written as

$$\begin{aligned} \mathbf{V}(\hat{\mathbf{X}}) &= [\mathbf{J}_{\hat{\alpha}} \ \mathbf{J}_Y] \begin{bmatrix} \mathbf{V}(\hat{\alpha}) & \mathbf{0} \\ \mathbf{0} & \mathbf{V}(\mathbf{Y}) \end{bmatrix} \begin{bmatrix} \mathbf{J}_{\hat{\alpha}}^T \\ \mathbf{J}_Y^T \end{bmatrix} = \\ &= \mathbf{J}_{\hat{\alpha}} \mathbf{V}(\hat{\alpha}) \mathbf{J}_{\hat{\alpha}}^T + \mathbf{J}_Y \mathbf{V}(\mathbf{Y}) \mathbf{J}_Y^T. \end{aligned} \quad (18)$$

Therefore, the covariance matrix of $\hat{\mathbf{X}}$ is made of two terms, one related to traceability and another accounting for the instrumental noise. The generic matrix element V_{ij} has the form

$$\begin{aligned} V_{ij} &= u(\hat{X}_i, \hat{X}_j) = \sum_{h,k} \frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} \frac{\partial \hat{X}_j}{\partial \hat{\alpha}_k} u(\hat{\alpha}_h, \hat{\alpha}_k) \\ &\quad + \delta_{ij} \left(\frac{\partial \hat{X}_i}{\partial Y_j} \right)^2 u^2(Y_j) \end{aligned} \quad (19)$$

where $\delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$ is the Kronecker symbol. As expected, the stimulus estimates are correlated by the common calibration coefficients. This result is analogous to that obtained in [8, Sec. 6] concerning the correlation introduced in mass comparisons by the common balance sensitivity. Quite in agreement with common sense, the instrumental noise contributes only to the estimate uncertainty and not to the covariance between any two estimates.

As concerns the derivatives in Eq. (19), we note that, according to the theory of implicit functions,

$$\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} = \left[\frac{\partial \hat{\alpha}_h}{\partial \hat{X}_i} \right]^{-1} \bigg|_{\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i)}, \quad (20)$$

and

$$\frac{\partial \hat{X}_i}{\partial Y_i} = \left[\frac{\partial Y_i}{\partial \hat{X}_i} \right]^{-1} \bigg|_{\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i)}. \quad (21)$$

In the case of a polynomial fit, one has in general

$$Y = \sum_{h=0}^{n-1} \alpha_h X^h. \quad (22)$$

and

$$\alpha_h = X^{-h} \left[Y - \sum_{k=0}^{n-1} \alpha_k X^k (1 - \delta_{hk}) \right] \quad (23)$$

Using Eq. (21) with Eq. (22) we obtain

$$\frac{\partial \hat{X}_i}{\partial Y_i} = \left[\sum_{k=1}^{n-1} k \hat{\alpha}_k \hat{X}_i^{k-1} \right]^{-1} \bigg|_{\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i)}. \quad (24)$$

The corresponding equation for the derivative with respect to $\hat{\alpha}_j$ is

$$\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} = \hat{X}_i^{2h} \left[-j Y_i \hat{X}_i^{j-1} - \sum_{k=0}^{n-1} (k-j) \hat{\alpha}_k \hat{X}_i^{k+j-1} \right]^{-1} \bigg|_{\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i)}, \quad (25)$$

which, after substitution of Eq. (22), becomes

$$\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} = -\hat{X}_i^h \left[\sum_{k=1}^{n-1} k \hat{\alpha}_k \hat{X}_i^{k-1} \right]^{-1} \bigg|_{\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i)}, \quad (26)$$

that is

$$\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} = -\hat{X}_i^h \frac{\partial \hat{X}_i}{\partial Y_i} \bigg|_{\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i)}. \quad (27)$$

This result can also be obtained directly from the theory of implicit functions. Equations (27) and (24) yield the sensitivity coefficients of the \hat{X}_i estimate with respect to the $\hat{\alpha}_j$ coefficients and to the Y_i response, respectively, and can easily be solved for any polynomial degree.

As already mentioned, an instrument is preferably used by taking the difference $\hat{X}_i - \hat{X}_j$ between two stimula corresponding to responses Y_i and Y_j . This difference may vary between near zero values, when the instrument is a comparator, or full-scale values, when the value X_j is the zero reading of the instrument. In order to evaluate the difference uncertainty, one can use Eq. (19). However, we follow a slightly different derivation. In matrix terms, the generic difference of any two estimates $\hat{X}_i - \hat{X}_j$ is modelled by $\hat{X}_i - \hat{X}_j = \mathbf{d}^T \hat{\mathbf{X}}$, where $\mathbf{d}_{1 \times q}^T = [0 \cdots 1_i \cdots -1_j \cdots 0]$. The resulting difference covariance is therefore

$$\mathbf{V} \left(\mathbf{d}^T \hat{\mathbf{X}} \right) = \mathbf{d}^T \mathbf{V} \left(\hat{\mathbf{X}} \right) \mathbf{d}, \quad (28)$$

or

$$\mathbf{V} \left(\mathbf{d}^T \hat{\mathbf{X}} \right) = \mathbf{d}^T \mathbf{J}_{\hat{\alpha}} \mathbf{V}(\hat{\alpha}) \mathbf{J}_{\hat{\alpha}}^T \mathbf{d} + \mathbf{d}^T \mathbf{J}_Y \mathbf{V}(Y) \mathbf{J}_Y^T \mathbf{d}. \quad (29)$$

By separately developing the two terms of the r.h.s of Eq. (29), we obtain

$$\mathbf{d}^T \mathbf{J}_{\hat{\alpha}} \mathbf{V}(\hat{\alpha}) \mathbf{J}_{\hat{\alpha}}^T \mathbf{d} = (\mathbf{J}_{i,\hat{\alpha}} - \mathbf{J}_{j,\hat{\alpha}})^T \mathbf{V}(\hat{\alpha}) (\mathbf{J}_{i,\hat{\alpha}} - \mathbf{J}_{j,\hat{\alpha}}), \quad (30)$$

where the row vector $\mathbf{J}_{i,\hat{\alpha}(1 \times n)} = \frac{\partial \hat{X}_i}{\partial \hat{\alpha}}$ is the i -th row of $\mathbf{J}_{\hat{\alpha}}$. Equation (30) yields

$$\begin{aligned} \mathbf{d}^T \mathbf{J}_{\hat{\alpha}} \mathbf{V}(\hat{\alpha}) \mathbf{J}_{\hat{\alpha}}^T \mathbf{d} = & \left. \frac{\partial \hat{X}_i}{\partial \hat{\alpha}} \mathbf{V}(\hat{\alpha}) \frac{\partial \hat{X}_i}{\partial \hat{\alpha}}^T \right|_{\hat{X}_i = g^{-1}(\hat{\alpha}, Y_i)} \\ & + \left. \frac{\partial \hat{X}_j}{\partial \hat{\alpha}} \mathbf{V}(\hat{\alpha}) \frac{\partial \hat{X}_j}{\partial \hat{\alpha}}^T \right|_{\hat{X}_j = g^{-1}(\hat{\alpha}, Y_j)} \\ & - 2 \left. \frac{\partial \hat{X}_i}{\partial \hat{\alpha}} \mathbf{V}(\hat{\alpha}) \frac{\partial \hat{X}_j}{\partial \hat{\alpha}}^T \right|_{\hat{X}_{i,j} = g^{-1}(\hat{\alpha}, Y_{i,j})}. \end{aligned} \quad (31)$$

This expression is written in the more familiar form

$$\mathbf{d}^T \mathbf{J}_{\hat{\alpha}} \mathbf{V}(\hat{\alpha}) \mathbf{J}_{\hat{\alpha}}^T \mathbf{d} = u_{\hat{\alpha}}^2(\hat{X}_i) + u_{\hat{\alpha}}^2(\hat{X}_j) - 2u_{\hat{\alpha}}(\hat{X}_i, \hat{X}_j). \quad (32)$$

Rearranging of the terms of Eq. (31) yields

$$\begin{aligned} & \mathbf{d}^T \mathbf{J}_{\hat{\alpha}} \mathbf{V}(\hat{\alpha}) \mathbf{J}_{\hat{\alpha}}^T \mathbf{d} = \\ & \sum_{h=0}^{n-1} \sum_{k=0}^{n-1} \left(\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} - \frac{\partial \hat{X}_j}{\partial \hat{\alpha}_h} \right) \left(\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_k} - \frac{\partial \hat{X}_j}{\partial \hat{\alpha}_k} \right) \bigg|_{\hat{X}_{i,j} = g^{-1}(\hat{\alpha}, Y_{i,j})} u(\hat{\alpha}_h, \hat{\alpha}_k), \end{aligned} \quad (33)$$

or

$$\begin{aligned} \mathbf{d}^T \mathbf{J}_{\hat{\alpha}} \mathbf{V}(\hat{\alpha}) \mathbf{J}_{\hat{\alpha}}^T \mathbf{d} = & \sum_{h=0}^{n-1} \left(\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} - \frac{\partial \hat{X}_j}{\partial \hat{\alpha}_h} \right)^2 \bigg|_{\hat{X}_{i,j}=g^{-1}(\hat{\alpha}, Y_{i,j})} u^2(\hat{\alpha}_h) \\ + 2 \sum_{h=0}^{n-2} \sum_{k=h+1}^{n-1} & \left(\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_h} - \frac{\partial \hat{X}_j}{\partial \hat{\alpha}_h} \right) \left(\frac{\partial \hat{X}_i}{\partial \hat{\alpha}_k} - \frac{\partial \hat{X}_j}{\partial \hat{\alpha}_k} \right) \bigg|_{\hat{X}_{i,j}=g^{-1}(\hat{\alpha}, Y_{i,j})} u(\hat{\alpha}_h, \hat{\alpha}_k). \end{aligned} \quad (34)$$

An analogous development for the second term in the r.h.s of Eq. (29) yields

$$\mathbf{d}^T \mathbf{J}_{\mathbf{Y}} \mathbf{V}(\mathbf{Y}) \mathbf{J}_{\mathbf{Y}}^T \mathbf{d} = \sum_{r=1}^q \left(\frac{\partial \hat{X}_i}{\partial Y_r} - \frac{\partial \hat{X}_j}{\partial Y_r} \right)^2 \bigg|_{\hat{X}_{i,j}=g^{-1}(\hat{\alpha}, Y_{i,j})} u^2(Y_r); \quad (35)$$

(the \mathbf{Y} covariance $\mathbf{V}(\mathbf{Y})$ being diagonal, the covariance terms vanish). The derivatives also vanish for any $r \neq i, j$, so that the only two remaining terms give

$$\mathbf{d}^T \mathbf{J}_{\mathbf{Y}} \mathbf{V}(\mathbf{Y}) \mathbf{J}_{\mathbf{Y}}^T \mathbf{d} = \left[\left(\frac{\partial \hat{X}_i}{\partial Y_i} \right)^2 + \left(\frac{\partial \hat{X}_j}{\partial Y_j} \right)^2 \right] \bigg|_{\hat{X}_{i,j}=g^{-1}(\hat{\alpha}, Y_{i,j})} u^2(Y_r). \quad (36)$$

It is worth noting that expressions (33) and (36) are quite general, and do not depend neither on the specific calibration function (3) chosen, nor on the assumption on the uncertainty of the X_i stimula used for calibration. Only Eq. (36) is based on assumption (7), but can easily be generalised, thus taking a form similar to that of Eq. (33).

4. The linear case

In the linear case one has

$$\hat{X}_i = -\frac{\hat{\alpha}_0}{\hat{\alpha}_1} + \frac{1}{\hat{\alpha}_1} Y_i. \quad (37)$$

In this case the derivatives of Eqs. (34) and (36) are easily calculated, and Eq. (29) becomes

$$u^2(\hat{X}_i - \hat{X}_j) = \frac{2u^2(Y_r) + (\hat{X}_i - \hat{X}_j)^2 u^2(\hat{\alpha}_1)}{\hat{\alpha}_1^2}. \quad (38)$$

The first contribution comes from the instrumental noise. The second comes from calibration, and thus accounts for traceability uncertainty. Its value is proportional to $(\hat{X}_i - \hat{X}_j)^2$, as conjectured. It is to be noted that the intercept uncertainty plays no role.

The relative uncertainty $w(\hat{X}_i - \hat{X}_j)$ takes the form

$$w(\hat{X}_i - \hat{X}_j) = \sqrt{w^2(Y_i - Y_j) + w^2(\hat{\alpha}_1)}. \quad (39)$$

In the case of linear fit, the covariance $\mathbf{V}(\hat{\mathbf{X}})$ can also be obtained directly from Eq. (18). By developing calculations, one has

$$\mathbf{V}(\hat{\mathbf{X}}) = \frac{1}{\hat{\alpha}_1^2} \left\{ [-\mathbf{e} - \mathbf{X}] \mathbf{V}(\hat{\alpha}) \begin{bmatrix} -\mathbf{e}^T \\ -\mathbf{X}^T \end{bmatrix} + u^2(Y_r) \mathbf{I} \right\}, \quad (40)$$

where $\mathbf{e}^T = [1 \dots 1]$. Explicit expressions for the covariance matrix elements are:

$$u^2(\hat{X}_i) = \frac{u^2(Y_r) + u^2(\hat{\alpha}_0) + \hat{X}_i^2 u^2(\hat{\alpha}_1) + 2\hat{X}_i u(\hat{\alpha}_0, \hat{\alpha}_1)}{\hat{\alpha}_1^2} \quad (41)$$

for variance and

$$u(\hat{X}_i, \hat{X}_j) = \frac{u^2(\hat{\alpha}_0) + \hat{X}_i \hat{X}_j u^2(\hat{\alpha}_1) + (\hat{X}_i + \hat{X}_j) u(\hat{\alpha}_0, \hat{\alpha}_1)}{\hat{\alpha}_1^2} \quad (42)$$

for covariance, from which Eq. (38) follows. Again, note that in the covariance term of Eq. (42), the only contribution comes from the coefficient uncertainty, as predicted in Eq. (19).

5. Higher-order fits

The expression for the difference uncertainty becomes more and more complicated as the degree of the fit increases. We report here the case of a second-degree polynomial

$$\begin{aligned} u^2(\hat{X}_i - \hat{X}_j) &= (\hat{X}_i - \hat{X}_j)^2 (B_i B_j)^{-2} \\ &\left\{ 4\hat{\alpha}_2^2 u^2(\hat{\alpha}_0) + \hat{\alpha}_1^2 u^2(\hat{\alpha}_1) + \left[\hat{\alpha}_1 (\hat{X}_i + \hat{X}_j) + 2\hat{\alpha}_2 \hat{X}_i \hat{X}_j \right]^2 u^2(\hat{\alpha}_2) + \right. \\ &\quad + 2\hat{\alpha}_1 \hat{\alpha}_2 u(\hat{\alpha}_0, \hat{\alpha}_1) + 2\hat{\alpha}_2 \left[\hat{\alpha}_1 (\hat{X}_i + \hat{X}_j) + 2\hat{\alpha}_2 \hat{X}_i \hat{X}_j \right] u(\hat{\alpha}_0, \hat{\alpha}_2) + \\ &\quad \left. + \hat{\alpha}_1 \left[\hat{\alpha}_1 (\hat{X}_i + \hat{X}_j) + 2\hat{\alpha}_2 \hat{X}_i \hat{X}_j \right] u(\hat{\alpha}_1, \hat{\alpha}_2) \right\} + \\ &\quad + (B_i^{-2} + B_j^{-2}) u^2(Y_r) \end{aligned} \quad (43)$$

where $B_{i,j} = \hat{\alpha}_1 + 2\hat{\alpha}_2\hat{X}_{i,j}$. Also in this case the uncertainty contributions due to calibration depend on the term $\hat{X}_i - \hat{X}_j$. The monomial representation (4) chosen for the polynomial expansion is known to lead to severe numerical difficulties [3], specially when the polynomial degree is higher than, say, 3 or 4. Yet, this representation is still widely adopted, an outstanding example being the reference function of the ITS90 [9]. Orthogonal polynomials have been proposed as an advantageous alternative [10, 11] and we think that their properties as concerns the in-use uncertainty deserves further investigation.

6. Comments

The derivatives in Eqs. (33) and (36) are the slopes of the curve (15), evaluated at the appropriate points. For $\hat{X}_i \rightarrow \hat{X}_j$ the calibration contribution given by Eq. (33) tends to zero and only the comparator noise term (36), tending to $\sqrt{2} \frac{\partial \hat{X}_i}{\partial Y_j} u(Y_j)$ remains. This behaviour holds independent of the curve function. In particular, in the case of fitting by a straight line, the calibration contribution is strictly proportional to $(\hat{X}_i - \hat{X}_j)^2$.

These results are quite in agreement with common sense. First, the different weight of the calibration contribution according to the different use of the instrument – as a comparator, thus resorting to an external standard, in case (1), or as a traceable standard, thus bearing all the uncertainty contribution arising from the calibration, in case (2) – is in agreement with everyday experience. Second, the noise contribution from the instrument has to be taken twice into account, in both cases. This result, apparent in the comparator mode, since two readings are needed, has sometimes been questioned for the direct reading mode, in which only one reading seems to be necessary. Actually, the preliminary “zeroing” of the instrument represents in any case a second reading, whose contribution to uncertainty cannot be neglected.

References

1. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, *Guide to the expression of uncertainty in measurement*, Geneva, 1993-1995.
2. I. H. Lira and W. Woeger, *Meas. Sci. Technol.* 8, 441 (1997).
3. M. G. Cox, A. B. Forbes, P. M. Harris and I. M. Smith, *The classification and solution of regression problems for calibration*, NPL Report CMSC 24/03, May 2003, 46 p.
4. W. Bich, in *Advanced Mathematical Tools in Metrology III*, (P. CIARLINI, M. G. COX, F. PAVESE and D. RICHTER Editors), Series on Advances

- in *Mathematics for Applied Sciences*, Vol. 45, World Scientific, Singapore (1997).
5. N. R. Draper, H. Smith, *Applied Regression Analysis*, John Wiley & Sons, New York (1966).
 6. A. B. Forbes, P. M. Harris, and I. M. Smith, in *Algorithms for Approximation IV*, (J. LEVESLEY, I. ANDERSON and J. C. MASON Editors), University of Huddersfield, 270 (2002).
 7. S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem*, SIAM, Philadelphia (1991).
 8. W. Bich, M. G. Cox and P. M. Harris, *Metrologia* 30, 495 (1993/94).
 9. H. Preston-Thomas, *Metrologia* 27/1, 3 (1990).
 10. D. R. White and P. Saunders, *Metrologia* 37, 285 (2000).
 11. D. R. White, *Metrologia* 38, 63 (2001).

AUTOMATIC DIFFERENTIATION AND ITS APPLICATION IN METROLOGY*

R. BOUDJEMAA, M. G. COX, A. B. FORBES, P. M. HARRIS

*Centre for Mathematics and Scientific Computing
National Physical Laboratory, Queens Road, Teddington, UK
Email: (redouane.boudjemaa, maurice.cox, alistair.forbes,
peter.harris)@npl.co.uk*

This paper describes automatic differentiation techniques and their use in metrology. Many models in metrology are nonlinear and the analysis of data using these models requires the calculation of the derivatives of the functions involved. While the rules for differentiation are straightforward to understand, their implementation by hand is often time consuming and error prone. The complexity of some problems makes the use of hand coded derivatives completely unrealistic. Automatic differentiation (AD) is a term used to describe numerical techniques for computing the derivatives of a function of one or more variables. The use of AD techniques potentially allows a much more efficient and accurate way of obtaining derivatives in an automated manner regardless of the problem complication. In this paper, we describe a number of these techniques and discuss their advantages and disadvantages.

1. Introduction

There is a wide variety of mathematical and scientific problems in which it is necessary to obtain accurate derivative evaluations. We recall that if $f(x)$ is a function of a variable x , the derivative

$$\frac{df}{dx}(x) \text{ or } f'(x)$$

is also a function of x and represents the slope of (the tangent to) f at x . Mathematically, the derivative is defined as a limit

$$\frac{df}{dx} = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}. \quad (1)$$

The fraction on the right represents the slope of the line passing through $(x, f(x))$ and nearby point $(x+h, f(x+h))$.

*Work partially funded under EU SoftTools_MetroNet Contract N. G6RT-CT-2001-05061

Using differentiation rules, the derivatives of complex functions can be expressed in terms of the derivatives of their simpler, component functions. However, even for only moderately complicated functions, the hand calculation of derivatives can lead to pages of tedious algebra with an increasing likelihood of a mistake entering into the computation.

If the function evaluation is encoded in a software component, it is natural to ask if it is possible to compute the derivatives automatically using the function evaluation component. Until recently, the standard method of evaluating derivatives numerically was to use finite differences, essentially evaluating the right-hand side of (1) with h a small, pre-assigned nonzero number. This approach generally gives an approximate value. In recent years, advances in computers and computer languages have allowed the development of a new method for obtaining accurate derivatives of any programmable function. The term *automatic differentiation*¹⁰ (AD) generally applies to techniques that produce, from a function evaluation software component, a computational scheme, also implemented in software, that calculates the derivatives. These techniques have evolved and are still evolving both in terms of their theoretical basis and, more markedly, in the software engineering aspects of their implementation. In a separate development, attention is now being paid to the *complex step method* which uses complex arithmetic to evaluate accurate derivatives.

In this paper we describe the main techniques for function differentiation, giving a summary of their advantages and disadvantages in terms of both accuracy and ease of implementation. The remainder of this paper is organized as follows. In section 2, we give a number of metrological application areas in which the calculation of derivatives is required. In sections 3, 4 and 5 we give descriptions of the main techniques for calculating derivatives and summarize their advantages and disadvantages. Example calculations are given in section 6. Our concluding remarks are presented in section 7.

2. Nonlinear problems in data analysis in metrology

In this section, we consider some of the main examples where derivative evaluation is required in metrology.

2.1. Nonlinear least squares approximation

In many calibration problems, the observation equation involving measurement y_i can be expressed as $y_i = \phi_i(\mathbf{a}) + \epsilon_i$, where ϕ_i is a function depending

on parameters $\mathbf{a} = (a_1, \dots, a_n)^T$ that specifies the behaviour of the instrument and ϵ_i represents the mismatch between the model specified by \mathbf{a} and the measurement y_i . Given a set of measurement data $\{y_i\}_1^m$, estimates of the calibration parameters \mathbf{a} can be determined by solving

$$\min_{\mathbf{a}} F(\mathbf{a}) = f_1^2(\mathbf{a}) + f_2^2(\mathbf{a}) + \dots + f_m^2(\mathbf{a}) = \sum_{i=1}^m f_i^2(\mathbf{a}), \quad (2)$$

where $f_i(\mathbf{a}) = y_i - \phi_i(\mathbf{a})$. Least squares techniques are appropriate for determining parameter estimates for a broad range of calibration problems⁷.

A common approach to solving least squares problems (2) is the Gauss-Newton algorithm. If \mathbf{a} is an estimate of the solution and J is the *Jacobian matrix* defined at \mathbf{a} by $J_{ij} = \partial f_i / \partial a_j$, then an updated estimate of the solution is $\mathbf{a} + \alpha \mathbf{p}$, where \mathbf{p} solves

$$\min_{\mathbf{p}} \|J\mathbf{p} + \mathbf{f}\|^2,$$

and α is a step length chosen to ensure that adequate progress to a solution is made. Starting with an appropriate initial estimate of \mathbf{a} , these steps are repeated until convergence criteria are met. The solution of a nonlinear least squares problem therefore requires the calculation of the derivative matrix J .

2.2. Uncertainty evaluation

If $U_{\mathbf{a}}$ is the uncertainty (covariance) matrix associated with a set of parameters \mathbf{a} and $f(\mathbf{a})$ is a function of \mathbf{a} then the standard uncertainty $u(f)$ associated with f is estimated from

$$u^2(f) = \mathbf{g}^T U_{\mathbf{a}} \mathbf{g}, \quad (3)$$

where \mathbf{g} is the gradient vector of f , i.e., $\mathbf{g} = (\partial f / \partial a_1, \dots, \partial f / \partial a_n)^T$. These partial derivatives are often referred to as *sensitivity coefficients*^{6,2}. If the uncertainty matrix is diagonal with $U_{ii} = u_i^2$, $U_{ij} = 0$, $i \neq j$, then (3) simplifies to the perhaps more familiar

$$u^2(f) = \left(\frac{\partial f}{\partial a_1} \right)^2 u_1^2 + \dots + \left(\frac{\partial f}{\partial a_n} \right)^2 u_n^2 = \sum_{j=1}^n \left(\frac{\partial f}{\partial a_j} \right)^2 u_j^2. \quad (4)$$

2.3. Maximum likelihood estimation and numerical optimization

Least squares approximation methods are a form of maximum likelihood estimation in which the most likely explanation of the data is sought. For

pre-assigned statistical models in which uncertainty in the measurement data is modelled by Gaussian distributions with a known uncertainty (covariance) matrix, the maximum likelihood corresponds to minimizing a sum of squares⁷. For more general models in which the uncertainty matrix is not known exactly, the maximum likelihood is attained by minimizing a function of the form

$$F(\mathbf{a}, \boldsymbol{\sigma}) = h(\boldsymbol{\sigma}) + \sum_{i=1}^m f_i^2(\mathbf{a}, \boldsymbol{\sigma}),$$

where $\boldsymbol{\sigma}$ are additional optimization parameters that specify the unknown input uncertainty matrix⁸. In this more general setting, F is no longer a sum of squares and the Gauss-Newton algorithm can no longer be applied (at least with the same success).

The Gauss-Newton algorithm is in fact derived from the more general Newton's algorithm for minimizing a function $F(\mathbf{a})$. In the Newton algorithm, the search direction \mathbf{p}_N is computed as the solution of $H\mathbf{p}_N = -\mathbf{g}$, where $\mathbf{g} = \nabla_{\mathbf{a}}F$ is the gradient vector of first partial derivatives, $g_j = \partial F / \partial a_j$, and $H = \nabla_{\mathbf{a}}^2 F$ is the *Hessian* matrix of second partial derivatives $H_{jk} = \partial^2 F / \partial a_j \partial a_k$.

3. Finite differences

3.1. Description

The main and simple idea of the method is to approximate the derivative (1) by evaluating f at two nearby points. Typical of finite difference formulae are the *forward difference* formula

$$f'(x) \approx \dot{f}_f = \frac{f(x+h) - f(x)}{h}$$

and the *central difference* formula

$$f'(x) \approx \dot{f}_c = \frac{f(x+h) - f(x-h)}{2h}.$$

Here h is a "step" selected in an appropriate way.

3.2. Advantages

The most important advantage of the finite difference approach is that it does not require access to the source code for the function evaluation component and therefore, from this point of view, it is an ideal method for library software.

3.3. Disadvantages

The method provides only approximate estimates. In ideal circumstances using central differences we still expect to lose one third of the available figures. For poorly scaled or ill-conditioned problems, the accuracy can be much worse. For metrology applications with data accurate to 1 part in 10^6 , there is a significant risk that finite differences implemented in IEEE arithmetic could introduce errors comparable to or larger than the measurement uncertainty. There are technical difficulties in balancing truncation errors with cancellation errors in the choice of step size h . The method is inefficient since for a problem of n parameters, it requires between $n + 1$ and $2n + 1$ function evaluations to calculate all n partial derivatives.

4. The complex step method

4.1. Description

The complex step method is similar to finite differences but uses complex arithmetic to provide accurate estimates. Current interest^{13,14} in the approach was generated by Squire and Trapp¹⁶ who revisited earlier work by Lyness and Moler¹². We recall that a complex number z is of the form $z = x + iy$ where x and y are real and $i = \sqrt{-1}$. All the arithmetic operations for real numbers also apply for complex numbers. Most real-valued functions $f(x)$ occurring in science also have complex-valued counterparts, and can be written in the form

$$f(z) = \Re f(z) + i\Im f(z),$$

where the real-valued functions $\Re f$ and $\Im f$ are known as the *real* and *imaginary parts*. The concept of derivative is defined by the complex version of (1) and from this we can also derive the Taylor expansion for a complex function. For x real and h real and small, we have

$$f(x + ih) = f(x) + ihf'(x) - \frac{h^2}{2}f''(x) - i\frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f''''(x) + \dots$$

Taking the imaginary part, we have

$$\Im f(x + ih) = hf'(x) - \frac{h^3}{3!}f'''(x) + \frac{h^5}{5!}f'''''(x) + \dots,$$

and we see that

$$f'(x) \doteq \dot{f}_z = \frac{\Im f(x + ih)}{h}, \quad (5)$$

with a truncation error of order h^2 . Unlike the use of a finite difference formula, h can be chosen to be *very* small (e.g. 10^{-100}) with no concern about the loss of significant figures through subtractive cancellation since no subtraction is involved.

The following Matlab instructions compares the CS estimate of the derivative of $\cos x$ with $\sin x$. The difference is calculated to be zero.

```
>> i=sqrt(-1); h = 1e-100; x=1; d = imag(sin(x+i*h)/h); d - cos(x)
```

We have implemented⁵, both in Matlab and Fortran 90, software components that use the CS method evaluate the derivatives of $f = f(\mathbf{x})$, $\mathbf{x} = (x_1, \dots, x_n)^T$, where f is calculated by a user-supplied component with \mathbf{x} declared as $n \times 1$ complex array.

4.2. Advantages

The complex step method can provide derivative information accurate to full machine precision. The method is easy to implement. In a language that supports complex arithmetic, the user can write a function evaluation component as normal, only declaring the relevant variables as complex, but care must be exercised with some intrinsic functions. There is no requirement to have access to source code, so long as the correct variable types have been used.

4.3. Disadvantages

As with standard finite differences, for problems involving a large number of variables, the method is inefficient. If there are n variables, then n function evaluations are required to calculate all n partial derivatives. Existing software written in real arithmetic has to be re-engineered to work with complex arithmetic. Subtractive cancellation can occur in special circumstances. Only first derivatives can be calculated.

5. Program differentiation techniques

Forward and reverse automatic differentiation (FAD, RAD) are two approaches aimed at “differentiating the program” that computes a function value⁹. They are accurate in the sense that they apply the rules of calculus in a repetitive way to an algorithmic specification of a function and, in exact arithmetic, will produce the exact answer. They rely on the fact that any program, no matter how complex it is, can be broken down into a

finite combination of elementary operators^{1,4,11} such as arithmetic operations (e.g., $+$, $-$) or elementary functions (e.g., $\sin x$, $\cos x$, e^x). Thus, the evaluation of a function $y = y(x_1, \dots, x_n)^T$ can be written as a sequence of m , say, operations

- 1 Initialise $t_j = x_j$, $j = 1, \dots, n$.
- 2 $t_q = \Psi_q(t_k, t_l)$, $q = n + 1, \dots, m$.
- 3 Set $y = t_m$.

Here Ψ_q is a unary, binary operation or elementary function. In FAD, the derivatives $\partial t_q / \partial x_j$ are calculated as the function evaluation progresses according to the chain rule:

$$\frac{\partial t_q}{\partial x_j} = \frac{\partial \Psi_q}{\partial t_k} \frac{\partial t_k}{\partial x_j} + \frac{\partial \Psi_q}{\partial t_l} \frac{\partial t_l}{\partial x_j}.$$

In a RAD approach, the function is evaluated in a forward sweep (as in FAD) and then in a reverse sweep the partial derivatives $\bar{t}_q = \partial y / \partial t_q$ of the output y with respect to the intermediate quantities t_q are assembled in reverse order, finishing with the required partial derivatives $\partial y / \partial x_j$. The reverse sweep is summarized as:

- 1 Initialize $\bar{t}_i = 0$, $i = 1, \dots, m - 1$, $\bar{t}_m = 1$.
- 2 For $q = m, m - 1, m - 2, \dots, n + 1$, if $t_q = \Psi_q(t_k, t_l)$,

$$\bar{t}_k := \bar{t}_k + \frac{\partial \Psi_q}{\partial t_k} \bar{t}_q, \quad \bar{t}_l := \bar{t}_l + \frac{\partial \Psi_q}{\partial t_l} \bar{t}_q.$$

- 3 For $j = 1, \dots, n$, $\partial y / \partial x_j = \bar{t}_j$.

The cost of computing the gradient of a function of n variables is between 2 and 6 times the cost of computing the function value itself for the RAD while for the FAD it is between n and $4n$.

Many of the derivative calculations involved in maximizing likelihoods reported in Cox *et al.*⁸ were performed using RAD.

5.1. Advantages

Both FAD and RAD provide mathematically accurate derivatives. RAD is particularly efficient for problems with large number of independent variables. The RAD memory efficiency requirements depend on the value m , i.e., the number of operations and elementary functions required to compute the function value. Higher derivatives can also be calculated.

5.2. Disadvantages

Both FAD and RAD require some effort to implement. For languages such as Fortran 90 that support operator overloading, FAD can be implemented by overloading the standard operations and elementary functions with operations on variable types that store derivative information in addition to function values. For RAD, information required for the reverse sweep has to be stored as the forward sweep progresses. The FAD and RAD algorithms require the modification of the declarative part of the source code to replace real variables with variables that store the required additional information. The rest of the code remains substantially the same.

6. Examples

6.1. $f(x) = x^{-1} \cos x^{-1}$

We have applied finite differences (FD), the complex step method (CS) and forward and reverse automatic differentiation (FAD, RAD) to provide estimates \hat{f} of the derivative f' of the function

$$f(x) = x^{-1} \cos x^{-1}, \quad f'(x) = \frac{x^{-1} \sin x^{-1} - \cos x^{-1}}{x^2}.$$

As x becomes smaller, the function and its derivative oscillate with increasing frequency and amplitude. We have calculated the derivatives at 100 points in the interval $[0.6, 6.0] \times 10^{-4}$ and computed the relative error $e = (f' - \hat{f})/|f'|$ for the four methods. In this interval the function attains values of the order of 10^4 and its derivative that of 10^{12} . We find that for CS, FAD and RAD, the relative error is no more than 5.0×10^{-16} , which compares favourably with the machine precision $\eta = 2.2 \times 10^{-16}$. By contrast, the relative error for the FD method ranges between 10^{-6} and 100.0.

6.2. Uncertainty in interferometric length measurements

The estimate of displacement d derived from the interferometric measurement of optical path length D is given by $d = D/n(t, p, v)$ where n is the bulk refractive index of the air along the light path and depends on temperature t , pressure p and humidity as measured by the partial vapour pressure v . In practice, the partial vapour pressure is often derived as a function $v(s)$ of the dewpoint temperature s . Given uncertainties associated with measurements D , t , p and s , the uncertainty associated with the geometric displacement d is calculated as in (4), section 2.2, and therefore involves

calculating the partial derivatives of $d = d(D, t, p, s)$. The revised Edlén formula³ for $n(t, p, v)$, for a fixed wavelength, is of the form

$$n(t, p, v) = 1 + c_1 p \left\{ \frac{1 + (c_2 - c_3 t)p}{1 + c_4 t} \right\} + c_5 v,$$

while that for $v(s)$ is of the form¹⁵

$$v(s) = \exp \{ c_6 / \bar{s} + c_7 + c_8 \bar{s} + c_9 \bar{s}^2 + c_{10} \log \bar{s} \}, \quad \bar{s} = s + c_{11},$$

where $\mathbf{c} = (c_1, \dots, c_{11})^T$ are fixed constants. (If the wavelength λ is regarded as a variable, then $c_1 = c_1(\lambda)$ and $c_5 = c_5(\lambda)$ are nonlinear functions of wavelength.) With these formulae, the evaluation of $d = d(D, t, p, s)$ has been encoded in software implementations in Matlab and Fortran 90, languages that support complex arithmetic. The evaluation of the sensitivity coefficients is therefore very simple using the complex step method.

7. Concluding remarks

In this paper, we have been concerned with methods for calculating the derivatives of functions. In particular, we have been interested in finite differences (FD), the complex step method (CS) and forward and reverse automatic differentiation (FAD and RAD). We have analyzed these methods in terms of accuracy, ease of implementation and efficiency with respect to speed and memory. We have found: 1) FD is the easiest to implement and use but is inaccurate, losing as much as half the available figures. FD is also inefficient for problems involving a large numbers of variables. 2) CS is easy to implement and use, accurate for first derivatives but can be inefficient for large problems. 3) FAD produces accurate first (and higher) derivatives, is straightforward to implement and use but can be inefficient for problems with a large number of independent variables. 4) RAD produces accurate first (and higher) derivatives and its memory requirements and computing times are determined by the complexity of the function, rather than the number of independent variables. It is quite straightforward to use but its implementation requires significant effort. More details can be found in Boudjemaa *et al.*⁵.

Acknowledgements

This work has been supported by the DTI National Measurement System Software Support for Metrology Programme (2001 – 2004). We are grateful to Prof. Bruce Christianson, University of Hertfordshire, for helpful discussions and comments on an earlier version of this work.

References

1. M. C. Bartholomew-Biggs, S. Brown, B. Christianson, and L. Dixon. Automatic differentiation of algorithms. *J. Comp. App. Math.*, 124:171–190, 2000.
2. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML. *Guide to the Expression of Uncertainty in Measurement*. Geneva, second edition, 1995.
3. K. P. Birch and M. J. Downs. Correction to the updated Edlen equation for the refractive index of air. *Metrologia*, 31:315–316, 1994.
4. C. Bischof, L. Roh, and A. Mauer. *ADIC – An extensible automatic differentiation tool for ANSI-C*. Argonne, IL. ANL/MCS-P626-1196.
5. R. Boudjemaa, M. G. Cox, A. B. Forbes, and P. M. Harris. Automatic differentiation and its applications to metrology. Technical Report CMSC 26/03, National Physical Laboratory, June 2003.
6. M. G. Cox, M. P. Dainton, and P. M. Harris. Software Support for Metrology Best Practice Guide No. 6: Uncertainty and Statistical Modelling. Technical report, National Physical Laboratory, Teddington, 2001.
7. M. G. Cox, A. B. Forbes, and P. M. Harris. Software Support for Metrology Best Practice Guide 4: Modelling Discrete Data. Technical report, National Physical Laboratory, Teddington, 2000.
8. M. G. Cox, A. B. Forbes, P. M. Harris, and J. Flowers. Least squares adjustment in the presence of discrepant data. *This volume*.
9. C. Faure. Automatic differentiation for adjoint code generation: An introduction to automatic adjoint code generation. Technical Report 3555, INRIA, 1998.
10. A. Griewank. On automatic differentiation. Technical report, Argonne National Laboratory, Argonne, IL, 1988.
11. A. Griewank and G. F. Corliss, editors. *Automatic Differentiation of Algorithms: Theory, Implementation and Applications*, Philadelphia, 1991. Society for Industrial and Applied Mathematics.
12. J. N. Lyness and C. B. Moler. Numerical differentiation of analytic functions. *SIAM J. Numer. Anal.*, 4:202–210, 1967.
13. J. R. R. A. Martins, I. M. Kroo, and J. J. Alonso. An automated method for sensitivity analysis using complex variables. In *38th Aerospace Sciences Meeting and Exhibit*, pages 10–13. American Institute of Aeronautics and Astronautics, Jan 2000.
14. J. R. R. A. Martins, P. Sturdza, and J. J. Alonso. The connection between the complex-step derivative approximation and algorithmic differentiation. In *39th Aerospace Sciences Meeting and Exhibit*. American Institute of Aeronautics and Astronautics, 2001.
15. D. Sontag. Important new values of the physical constants of 1986, vapour pressure formulations based on ITS-90, and psychrometer formulae. *Z. Meteorol.*, 40(5):340–344, 1990.
16. W. Squire and G. Trapp. Using complex variables to estimate derivatives of real functions. *SIAM Rev.*, 40:110–112, 1998.

USAGE OF NON-CENTRAL PROBABILITY DISTRIBUTIONS FOR DATA ANALYSIS IN METROLOGY

A. CHUNOVKINA

*D.I. Mendeleyev Institute for Metrology, Moskovsky pr. 19,
St. Petersburg, 190005, Russia
E-mail: A.G.Chunovkina@vniim.ru*

The methods of statistical hypotheses testing are widely used for the data analysis in metrology. As a rule they are based on normal distribution, t -distribution, χ^2 -distribution and F -distribution. The criterion of hypotheses testing is usually characterized only by the level of significance. Such characteristic as a criterion power is not actually used in metrological practice. The paper discusses the problems of using the corresponding non-central distributions for the evaluation of the criterion power as well as actual level of significance in the present of systematic biases in measurement results. The examples of testing the measurand value, the difference between measurement results, the consistency of the data and the relations' model are considered.

1. Introduction

Statistical methods for testing hypotheses are widely used in metrology in case, when a certain decision should be made on the basis of measurement results obtained. A few examples are given below:

- Testing the deviation of the value of the parameter from its specified standard value (testing reference values of standard materials),
- Testing the significance of systematic differences between measurement results obtained by various measurement procedures,
- Testing the consistency of experimental data with a model suggested (model of functional relations),
- Testing the homogeneity of the data, received from various sources (testing the difference between the means and the variances),
- Others

A conventional Neyman and Pearson approach for testing hypotheses is applied [1]. If a measurement result, x , is in the region of the values, which are consistent with the null hypothesis, the null hypothesis H_0 is accepted. Otherwise, if the measurement result is in critical region Ω_K - H_0 is rejected. $\Omega_K = \Omega \setminus \Omega_A$, where Ω - is a region of all possible measurement results and Ω_A is a region of values consistent with the null hypothesis. As a rule, percentiles for normal, χ^2 , t and \mathcal{F} distributions are usually used as critical

values at the determination of the region of values consistent with the H_0 . The critical region is determined so as that the probability of incorrect rejection of the null hypothesis (type I error probability or a criterion level of significance) would be small enough (usually accepted values are 0,05 or 0,01):

$$P(x \in \Omega_K | H_0) = \alpha. \quad (1)$$

Here the Neyman and Pearson interpretation of the level of significance (significance level) should not be mixed with the significance level of observed values (measurement results). The latter interpretation will be also used in this paper in 2.5.

In using statistical methods for testing hypotheses in metrology, at least two aspects of their correct application should be considered:

1. The acceptance of the null hypothesis, in fact, means an acceptance of the related model. It can be a specified value of a parameter or a relation model between variables and other. The model is chosen as adequate to the data. Then it can be applied to data processing or to measurement design. However, strictly speaking, the question on the model accuracy remains open, as the acceptance of a model does not yet mean its "trueness". Different models can be accepted on the basis of the same criterion and on the basis of the same measurement results. Their subsequent application leads to the results, and the deviations between them can exceed the associated uncertainties assigned to these results. This is easily explained by the fact that the uncertainty of the measurement results is evaluated within the frame of the model approved. Since it is difficult to characterize quantitatively the uncertainty of the model accepted, it is important to search for another way to characterize the validity of the model used. One of the possible ways is to estimate the probability of incorrect approval of the null hypothesis related to the model considered. In other words the criterion power should be estimated.
2. The application of statistical methods of testing hypotheses, as well as of statistical estimations of parameters, in metrology requires the correct handling and treatment of the systematic errors of measurement results. The systematic biases in measurement results should be taken into account when level of significance and criterion power are actually estimated, as well as when the judgment on the measurement results is made.

The paper considers these aspects applied to some concrete examples.

2. Example A. Testing the null hypothesis about equality of the measurand value and a specified standard value

The following task is considered. Testing the null hypothesis $H_0 : a = a_0$, where a – actual measurand value, a_0 – specified standard value, is considered. It's assumed that the measurement results X_1, \dots, X_n come from normal distribution with parameters $N(a, \sigma^2)$. Under H_0 the following test statistic

$$T = \frac{(\bar{X} - a)\sqrt{n}}{S}, \quad (2)$$

where $\bar{X} = \frac{1}{n} \sum X_i$, $S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$, is distributed as t -Student distribution with $n-1$ degrees of freedom. Hence, the region of values, which are consistent with H_0 , is defined by the equation:

$$|t| \leq t_{0.95}(n-1), \text{ where } t = \frac{(\bar{x} - a)\sqrt{n}}{S}. \quad (3)$$

According to the above, the following questions should be answered:

1. What is the probability to reject H_0 , when it is not true? How should the alternative hypothesis H_1 be formulated?
2. How many repeated measurements $\{n\}$ should be carried out to reject H_0 with the given probability (power) when the actual deviation of the measurand value from the specified value is more than Δa ?
3. What are the actual type I and type II errors probabilities when the conventional well-known criterion with the significance levels assigned in advance are used in the presence of systematic biases in measurement results?
4. Should any correction be made according to the test results? Or should the original experimental deviation from the specified value be taken used for the uncertainty evaluation?

2.1. Estimation of the criterion power

The answer to the first previous question requires the formulation of the alternative hypothesis H_1 . To do this, we should use some additional information, which is usually available in practice. For example, when testing the deviation from the standard value (in quality control inspection or other), the

limit of the permitted deviations is usually specified. Denote the limit of the neglected deviations as Δa_0 . Then the alternative can be formulated as follows:

$$H_1 : |a - a_0| \geq \Delta a_0 \quad (4)$$

The maximum Type II error probability (the risk of incorrect concluding that H_0 is true) is realized for actual measurand values equal to $a = a_0 \pm \Delta a_0$:

$$\beta = \max_a P(H_0 | H_1) = P(|t| \leq t_{0,05}(n-1) | a = a_0 \pm \Delta a_0) \quad (5)$$

Consequently the criterion power equals $1 - \beta$

Under H_1 ($a = a_0 + \Delta a_0$ or $a = a_0 - \Delta a_0$) the test statistics $T = \frac{(\bar{X} - a_0)\sqrt{n}}{S}$ is distributed as non-central t-distribution with parameter of non-centrality $c = \frac{\Delta a_0 \sqrt{n}}{\sigma}$.

To use the standard tables for criterion power values it is necessary to specify: n – degree of freedom, α – level of significance, c – parameter of non-centrality.

The increase of the above parameters results in the increase of the criterion power. First of all, it concerns the dependence upon the parameter of non-centrality. Coming back to the example considered, it should be stressed that, in the case given, the parameter of non-centrality itself is an increasing function of the number of repeated observations.

2.2. The choice of a number of the repeated observations

The standard tables for the criterion power values can be also used for a rational choice of a number of repeated observations n in case when level of significance and power of the criterion are specified. The alternative hypothesis H_1 is determined as above: $H_1 : |a - a_0| \geq \Delta a_0$.

The procedure of choosing n is valid when a random dispersion of the measurement results dominates their systematic biases. With the increasing of the number of repeated observations, the systematic and random effects become comparable. Further increasing the number of observations does not lead to an increase of criterion power, as it could be expected.

2.3. Testing the systematic differences between measurement results

It is important to outline special cases, where the application of the t -criterion is correct:

1. testing the changes of the parameter under control. Here, the same measurement procedure is used, so systematic biases do not influence the differences between measurement results and should not be taken into account. For example, this task occurs when the stability of some process is under control. In the given case, H_0 relates to the assumption that the parameter value remains the same and consequently the differences between measurement results are caused only by the reproducibility of the measurement procedure used.
2. testing the systematic difference between measurement results obtained by different measurement procedures. Here it's known that the measurand value is the same for two different measurements. In the given case, H_0 relates to the assumption that there is no systematic bias between the two measurement results consequently the distinction between them is caused only by their random dispersion.

The null hypothesis is formulated as follows: $H_0 : a_1 = a_2$, where a_1, a_2 – the actual measurand values for the two measurements. (In a more general case, testing the null hypothesis $H_0: a_1 = a_2 = \dots = a_N$ is considered and F -test statistic is used). Under H_0 the difference between two measurement results should satisfy the following equation:

$$|\bar{x}_1 - \bar{x}_2| \leq t_\alpha (n_1 + n_2 - 2) \sqrt{\frac{n_1 + n_2}{n_1 n_2 (n_1 + n_2 - 2)}} \sqrt{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}$$

$$\bar{x}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} x_{1j} \quad S_1^2 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2 \quad (6)$$

$$\bar{x}_2 = \frac{1}{n_2} \sum_{j=1}^{n_2} x_{2j} \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2$$

where $t_\alpha (n_1 + n_2 - 2)$ – percentile of t -distribution with $n_1 + n_2 - 2$ degrees of freedom and a level of significance α .

For calculation of the criterion power, a non-central t -distribution with $n_1 + n_2 - 2$ degrees of freedom is used, with a parameter of non-centrality

$c = \frac{\Delta a_0}{\sigma} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$, where Δa_0 is the limit of the permitted differences as determined above.

2.4. The way of the handling systematic biases in measurement result

When the conventional well-known criteria are used in metrological practice in the presence of systematic biases in measurement results (which are not taken into account) the actual type I error probabilities differ from their expected values. Generally, systematic biases in measurement results lead to an increasing of the level of significance and to a decrease of the criterion power, comparing with their values in absence of systematic biases. It is obvious, that the null hypothesis and the alternative one cannot be discriminated when Δa_0 is less than the level of the systematic bias in the measurement results or of the difference of systematic biases depending on the task considered. It is suggested to take into account the systematic biases for the determination of the alternative hypothesis by putting $\Delta a_0 = \theta$, where θ is the limits of the systematic biases:

$$\begin{aligned} H_0 : \Delta a_0 &= 0 \\ H_1 : \Delta a_0 &= \theta \end{aligned} \quad (7)$$

Under the H_0 hypothesis, the differences between the expectation of the measurement results and the specified standard value can lay within the limits $(-\theta, \theta)$. So the level of significance (type I error probability) is suggested to be characterized by the average value within these limits:

$$\tilde{\alpha} = \frac{1}{2\theta} \int_{-\theta}^{\theta} P(|t| > t_{\alpha}(n-1) | a = a_0 + x) dx \quad (8)$$

Here, the uniform distribution of systematic biases is assumed as in Bayesian approach for testing hypotheses. To calculate $P(|t| > t_{\alpha}(n-1) | a = a_0 + x)$ the standard tables for the t -criterion power values

(based on non-central distribution with $c = \frac{x}{\sigma} \sqrt{n}$) are used. Hence, under H_1 , the test statistics considered above has a non-central t -distribution with parameter of non-centrality equal to $c = \frac{\theta}{\sigma} \sqrt{n}$:

$$\beta = P(|t| \leq t_{\alpha}(n-1) | a = a_0 + \theta)$$

If the procedure for the determination of the alternative hypothesis does not satisfy the user requirements, a more precise measurement procedure should be used for testing the hypotheses.

2.5. *Has the original values to be retained or the measurement uncertainty has to be increased in accordance with the test results?*

Let us consider the following situation. A specified standard value a_0 is used for the subsequent calculations and, consequently, it influences the accuracy of the final result. A natural question arises: is it necessary to take into account the experimental deviations of the measurement result from the specified value when estimating the quality of final results or not? If yes, how should it be made? As a matter of fact, it is a question about the "validity" ("trueness") of the model, which is accepted by testing the related hypotheses.

Within the frame of the approach considered (Neyman and Pearson approach for testing hypotheses), it is difficult to answer the above question. According to this approach the null hypothesis is accepted or rejected depending on whether the measurement result is in the critical area or not. Therefore, the actual magnitude of the deviation from standard value is not taken into account. To account for the deviation magnitude we need another interpretation of the significance level. Here it's interpreted as significance level of the data received (opposite to the above interpretation of criterion significance level) [2]. Here significance level characterizes the consistency (agreement) of the particular data obtained with the hypothesis tested. Note, in the given interpretation type I and type II errors probabilities are not used because no alternatives are considered. So the following procedure is suggested. Under H_0 the probability

of the event that test statistics $T = \frac{(\bar{X} - a_0)\sqrt{n}}{S}$ exceeds the actually observed

value, should be calculated: $\pi = P\left(T > \frac{\bar{x} - a_0}{S}\sqrt{n} \mid H_0\right)$.

Depending on the π values the following conventional procedure is applied:

π	<u>Interpretations</u>
$\pi > 0,10$	Strong argument for $H_0 \Rightarrow$ no correction
$\pi = 0,05$	Some doubts in $H_0 \Rightarrow$ further analysis
$\pi = 0,02$	Strong argument against $H_0 \Rightarrow$ correction
$\pi = 0,01$	H_0 is rejected

3. Example B. Testing the hypothesis about the model between variables

For testing the hypothesis about a model, the following test statistics is usually used (for simplicity a linear relations is considered):

$$v^2 = \frac{N(n-1)n \sum_{i=1}^N (\bar{y}_i - \hat{y}_i)^2}{(N-2) \sum_{i=1}^N \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2} \quad (9)$$

where $\{x_i, y_{ij}\}_{i=1, \dots, N, j=1, \dots, n}$ are experimental data

$$\begin{aligned} \bar{x} &= \frac{1}{N} \sum_{i=1}^N x_i & \hat{a}_0 &= \bar{y} = \frac{1}{N} \sum_{i=1}^N \bar{y}_i \\ \hat{y}_i &= \hat{a}_0 + \hat{b}(x_i - \bar{x}), & \hat{b} &= \frac{\sum \bar{y}_i (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \\ \bar{y}_i &= \frac{1}{n} \sum_{j=1}^n y_{ij} \end{aligned}$$

It is assumed that data pertain to normal distributions with expectations equal to measurand values and variances σ_x^2 , σ_y^2 . The case of controlled variable is considered so the least square estimation of linear relations is valid [3], [4], [5]. If

$$v^2 \leq F_{0.95}(\nu_1, \nu_2), \quad (10)$$

then the null hypothesis for linear relations is accepted. Here $F_{0.95}(\nu_1, \nu_2)$ is the 95% percentile of the *F-distribution*, with $\nu_1 = N-2$, $\nu_2 = N(n-1)$ degree of freedom.

The application of the *F-criterion* is valid when the independent variable x is known without errors. In practice, there are always errors also in the independent variable. Hence, the application of the *F-criterion*, strictly speaking, is incorrect. The test statistics Eq. (9) is distributed as a non-central *F*-distribution with parameter $a = b(N-1) \frac{\sigma_x^2}{\sigma_y^2}$ and degrees of freedom $(N-2, N(n-1))$.

The actual type I error probability for the criterion considered can be estimated using standard tables for *F-criterion* power values. To use these tables

it is necessary to determine a parameter $\varphi = \sqrt{\frac{a}{v_1 + 1}}$ and degrees of freedom.

For the case considered φ equals to $\varphi = \sqrt{bn} \frac{\sigma_x}{\sigma_y}$. Note that the increase of the number of repeated observation results in an increasing the probability of rejection of null hypothesis when it is true (type I error probability). For example, for $\varphi = 1$, $N = 4$ the actual type I error probability is in a range from 0,2 to 0,3 (depending on n) while the assigned value is 0,05.

4. Conclusions

The general problem of application of statistical methods for testing hypotheses in metrology has been considered on concrete examples. The following questions have been outlined:

1. Interpretation of judgments based on the results of testing hypotheses. Characterization of the quality of the model related to the hypothesis tested.
2. Account for the systematic biases in measurement results when formulating an alternative hypothesis
3. Using non-central distribution with:
 - estimate of the criterion power
 - estimate of actual level of significance when conventional test statistics are used in the presence of systematic biases.

References

1. Lehmann E.L. Testing statistical hypotheses, John Wiley, New York, 1959
2. Handbook of applicable mathematics, vol.VI: Statistics, part A, John Wiley & Sons Ltd
3. M.G.Kendall , A.Stuart The Advanced Theory of Statistics, v.2 Inference and relationship, Charles Griffin & Company Ltd, London
4. J.Berkson (1950) Are there two regressions? J.Amer.Statist.Ass. 45, 164
5. D.V.Lindley (1953) Estimation of functional relationship, Biometrika 40, 47

IMPLEMENTATION OF A GENERAL LEAST SQUARES METHOD IN MASS MEASUREMENTS*

J. HALD AND L. NIELSEN

*Danish Fundamental Metrology
B307 Matematiktorvet
DK-2800 Kgs. Lyngby, Denmark
E-mail: jha@dfm.dtu.dk*

We present a new software product (*DFM Calibration of Weights*) developed at Danish Fundamental Metrology (DFM), which has been implemented in mass measurements. This product is used for designing the mass measurements, acquiring measurement data, data analysis, and automatic generation of calibration certificates. Here we will focus on the data analysis, which employs a general method of least squares to calculate mass estimates and the corresponding uncertainties. This method provides a major simplification in the uncertainty calculations in mass measurements, and it allows a complete analysis of all measurement data in a single step. In addition, we present some of the techniques used for validation of the new method of analysis.

1. Introduction

The primary mass standard in Denmark is the national kilogram prototype no. 48. It is made of the same platinum-iridium alloy (90% Pt – 10% Ir) as the international kilogram prototype. Traceability to the international prototype, which is placed at the Bureau International des Poids et Mesures (BIPM) in Paris, is obtained by calibration of the Danish prototype against the prototype copies maintained at BIPM. The prototype no. 48 was cast in 1937-38 in France, and it was delivered by BIPM to Denmark in 1950. The Danish prototype is kept at DFM in a controlled environment, and it is used in calibration of DFM's stainless steel weights, which are in turn used as reference weights in client calibrations.

The process of mass calibration is quite complex, and the data acquisition and the analysis of the weighing results are most easily done by the use of computers. In the following we present the new software implemented for mass calibrations at DFM together with an introduction to the method for data analysis and some of the validation results.

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

2. The principle of subdivision and multiplication

We will describe the subdivision procedure by considering an example where a set of weights (1kg, 500g, 200g, 200g*, 100g, 100g*) with unknown masses is calibrated. First we extend the weight set by adding a reference weight (R1kg) with known mass. A measurement sequence is then established, where each line in the sequence provides an apparent mass difference of two weight combinations of the same nominal mass. A mass comparator is used for these mass measurements. The mass comparator can measure the mass difference between two weights of the same nominal mass with a very high resolution. Different mass comparators are often used in different ranges of mass values. A possible sequence of measurements is illustrated in the following equation, where the measured mass differences are written in matrix form:

$$\begin{pmatrix} \Delta I_1 \\ \Delta I_2 \\ \Delta I_3 \\ \Delta I_4 \\ \Delta I_5 \\ \Delta I_6 \\ \Delta I_7 \\ \Delta I_8 \\ \Delta I_9 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & -1 & 0 & -1 \\ 1 & 0 & -1 & -1 & -1 & -1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 & -1 \\ 0 & 0 & 1 & -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} m_{\text{R1kg}} \\ m_{\text{1kg}} \\ m_{\text{500g}} \\ m_{\text{200g}} \\ m_{\text{200g}^*} \\ m_{\text{100g}} \\ m_{\text{100g}^*} \end{pmatrix}, \text{ or } \Delta \mathbf{I} = \mathbf{X} \mathbf{m} \quad (1)$$

The design matrix \mathbf{X} on the right specifies the weights that enter in each comparison (each line in \mathbf{X}). The vector \mathbf{m} consists of the apparent masses of the weights (with only m_{R1kg} known in advance). The vector $\Delta \mathbf{I}$ holds the apparent mass differences calculated from the readings of the mass comparator. Each comparison is typically repeated 12 times or more in order to reduce the statistical uncertainty and to ensure repeatability in each measurement.

Eq. (1) is an over-determined set of linear equations, which do not have a solution because of the imperfect measurement of $\Delta \mathbf{I}$. However, by a linear least squares fit in which the known value of the reference mass is used as a constraint [1,2], it is possible to derive the 'best' \mathbf{m} -vector, which will minimize the difference between the left- and right-hand side of Eq. (1) in the l_2 norm.

It is of course possible to extend the design matrix to include weights of even smaller masses or weights with masses above 1 kg. However, it is often convenient to use separate design matrices for each decade of nominal mass values. This is normally accomplished by introducing combinations of weights from another decade as a single 'sum-weight'. As an example, weights of nominal masses 50g, 20g, 20g* and 10g can be grouped as an S100g weight and

entered into the 1 kg –100 g decade. Analysis of the 1 kg –100 g decade gives the mass value for the S100g sum, which will subsequently be used as the reference value in the 100 g –10 g decade.

3. The real measurement model

In the real world, however, the volumes and shapes of the weight combinations are not identical, and the measurements need correction for air buoyancy as well as for the gravity gradient. These effects are included in the following weighing equation for the i 'th line in the measurement sequence:

$$f_q \Delta I_i = \sum_j x_{ij} \left(1 + \frac{g'}{g} z_{i,j} \right) (m_j - a_i V_j) \quad (2)$$

Here x_{ij} are the elements of the design matrix X , g is the gravity acceleration at a certain reference height, and g' is the vertical gradient of g . $z_{i,j}$ is the vertical distance between the reference height and the center of mass position (CM position) of weight j . $z_{i,j}$ depends on i , because the weight j may be stacked on top of other weights in design line i . V_j is the volume of weight j , and f_q is a scale factor correction for the comparator readings. The density of air (a_i) is calculated from measurements of the temperature (t_i), pressure (p_i), dew-point temperature (td_i) and CO₂ contents ($xco_{2,i}$):

$$a_i = h(t_i, p_i, td_i, xco_{2,i}) \quad (3)$$

The formula h is the 1981/91 equation for the determination of the density of moist air recommended by the Comité International des Poids et Mesures (CIPM) [3,4]. This equation is not perfect, and an uncertainty is associated with this formula. Furthermore, each of the measurement instruments has a calibration curve described by a set of parameters with associated uncertainties and correlation coefficients. Thus, a very large number of measurement data and calibration parameters enter into the calculation of the unknown masses, and a complete uncertainty calculation according to the *ISO Guide to the Expression of Uncertainty in Measurement* (GUM) becomes quite complex. The traditional method, which is used at a number of National Metrology Institutes, is based on the least squares solution presented in section 2 together with a number of approximations and assumptions in order to apply the GUM method of uncertainty calculation. Furthermore, each decade of mass values and the

Table 1. A list of the different parameters (ζ and β) and constraints (f), which enters in a typical calibration of a weight set with 43 weights. The conventional mass of a weight is defined by OIML [7]. CM: center of mass.

ζ	β	f
96 Weighing results, ΔI_i	48 Mass values, m_j	92 Weighing equations, air density, instr. calib.
43 CM positions, z_j	48 Conventional mass values	2 Equations for reference mass values
43 Weight heights	43 Weight densities (or volumes)	5 Equations for sum weights
43 Weight volumes, V_j (or densities)	43 CM info	48 Equations for conv. masses
43 Thermal expansion coefficients		43 Equations for CM info
43 Mass values (only used for reference weights)		43 Equations for density (or volume)
18 Calibration parameters, e.g. f_q		
7 Temperatures for thermal expansion		
1 Correction factor in air density equation		
337 parameters	182 parameters	237 constraints

corresponding part of the design matrix are analyzed separately. The results for e.g. sum weights in one decade are then used as the “reference mass” in the next decade. This method was previously used at DFM.

4. The general least squares solution

In *DFM Calibration of Weights* we have used a general least squares method for the evaluation of mass estimates and uncertainties. We describe all input parameters, for which prior estimates and associated uncertainties are available, as elements of a vector z . The elements in z include the comparator readings in ΔI , information about weight volumes, heights and thermal expansion coefficients, and calibration parameters of the measurement instruments. The covariance matrix Σ describes the known uncertainties and possible correlations

for the input parameters \mathbf{z} . A vector $\boldsymbol{\beta}$ holds the parameters for which no prior knowledge is available, i.e. the quantities that we wish to determine (primarily masses and conventional masses [7]). An improved estimate for \mathbf{z} (named $\boldsymbol{\zeta}$) as well as an estimate for $\boldsymbol{\beta}$ is obtained by minimizing the weighted chi-square function with respect to $\boldsymbol{\zeta}$:

$$\chi^2(\boldsymbol{\zeta}; \mathbf{z}) = (\mathbf{z} - \boldsymbol{\zeta})^T \Sigma^{-1} (\mathbf{z} - \boldsymbol{\zeta}) \quad (4)$$

χ^2 is minimized under the constraints induced by the weighing equations, the air density formula, instrument calibrations etc. – all represented as a vector function \mathbf{f} fulfilling:

$$\mathbf{f}(\boldsymbol{\beta}, \boldsymbol{\zeta}) = 0 \quad (5)$$

A simple method to solve this problem is presented in [5,6], where (improved) output estimates $\boldsymbol{\zeta}$ and $\boldsymbol{\beta}$ as well as associated uncertainties and correlations are obtained as the output. If \mathbf{f} is a nonlinear function the solution is obtained by iterative calculations.

5. DFM Calibration of Weights

The *DFM Calibration of Weights* software consists of two modules. One module is based on a Microsoft Excel template with Visual Basic for Applications code. This module is used to specify the weights that enter into the calibration and to construct the weighing design X . This module is also used to automatically generate the calibration certificate as a Microsoft Word document based on the analyzed calibration data. Finally, the Excel module automatically updates the calibration database for DFM's own weights and for the client weights. The other module is written in LabWindows (C code) and it controls the weighing sequence and data acquisition from mass comparators and instruments for monitoring the climate. The LabWindows module also includes the algorithm for data analysis using the approach described in section 4.

As an example, we consider a typical calibration involving 43 mass standards, which are calibrated using 3 different mass comparators. The quantities collected in $\boldsymbol{\zeta}$ and $\boldsymbol{\beta}$ are described in Table 1, and we note that a total of 519 parameters are involved in the calculation. The 237 equations defining the constraints are also listed in Table 1. Note that not all parameters and equations from Table 1 are included in the simplified description in section 3 and 4. The analysis converges in 4 iterations, which takes about 2 minutes on a 500 MHz Pentium PC. The uncertainties from the calculation itself (e.g.

machine precision limitations on numerical derivatives, matrix operations etc.) on the derived mass values are all below 10^{-7} relative to the calculated mass uncertainties.

The advantage of the new method is that all decades in the calibration are analyzed in one step, and the consequences of changing one or more input parameters or uncertainties can be analyzed quickly. Furthermore, correlations between input parameters (typically the mass and volume of the reference weights) are easily included and no approximations of the involved equations are required.

6. Validation of *DFM Calibration of Weights*

DFM is accredited according to ISO 17025, and we must therefore fulfill the requirements on software validation given in this standard. *DFM Calibration of Weights* is validated using the Nordtest method and template for software validation [8]. DFM participated in the development of this validation method, and it is designed to comply with both the ISO 17025 and ISO 9001 requirements.

The Nordtest method is based on the software life cycle model, and it can be used both as a development tool and a validation tool. In Fig. 1 we show the different phases in the life cycle. The most important part of the validation is the testing part where the software is tested against the specified requirements. In the following we focus on the tests that ensure that *DFM Calibration of Weights* performs as required.

Careful inspection of the source code and dynamic testing was applied during the development phase. Such structural testing provides confidence that each module or function works correctly, but more general tests are required to see how the individual modules interact and to test the integration of all software modules with the hardware and peripheral units (i.e. mass comparators and environmental instruments).

To achieve correct calibration results, we must ensure that the data are transferred correctly from the different peripheral instruments to the computer and into the electronic files where raw data are stored. This is ensured by direct observation of values displayed on the instruments and by comparison to the values stored in the files. Furthermore, correct analysis of raw data must be checked, and finally it is checked that the analyzed data are transferred correctly into the final calibration certificate. The validation of the data analysis is difficult, since the employed method for the analysis is quite complex. Typically validations of calculations involve parallel testing, where manual calculations or

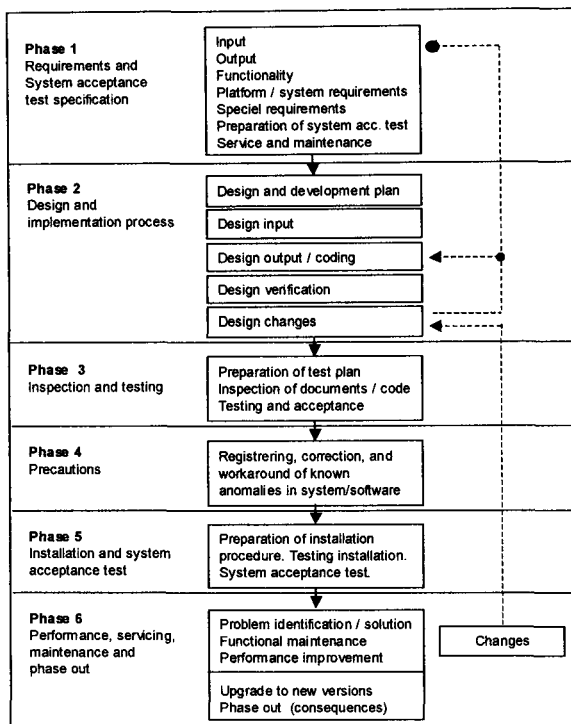


Fig 1. Schematic overview of the software life cycle model used in the Nordtest validation method.

alternative (but equivalent) software solutions are used to reproduce the results. Instead of direct parallel testing, we have used the following tests to ensure correct data analysis:

- Comparison of mass values obtained with new software to previous mass values for DFM weights
- Analysis of old raw data using the new software
- Analysis of a constructed test data set

In addition, we continuously monitor the calibration results by including a few of DFM's own weights as check standards whenever client weights are calibrated. By comparing the derived masses for the check standards with their calibration history, we can detect with a high probability if an error has occurred. In addition, the value for the minimized χ^2 value provides information on the consistency of the measurement results [5,6].

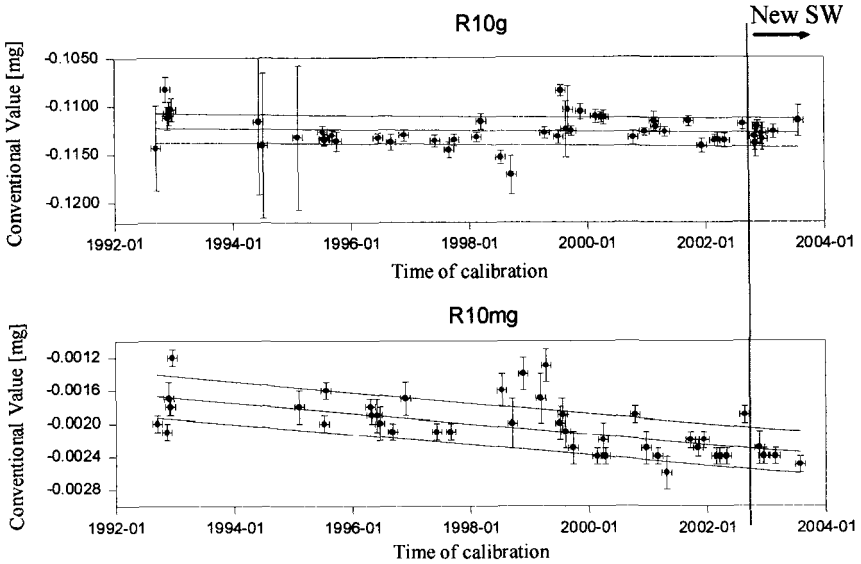


Fig.2. Calibration history of two DFM weights. The vertical scale is the difference between the conventional mass and the nominal mass. The new software was implemented in September 2002.

In Fig. 2 we show the calibration history of two weights from DFM's weight set. The error bars on the measurements are the standard uncertainties (one standard deviation) calculated in the data analysis. This uncertainty includes the contribution from the uncertainty of the 1 kg reference weight. However, this contribution is only about 10^{-8} in relative units, and for the weights in Fig. 2 this contribution can be neglected. Thus, the error bars in Fig. 2 reflect the uncertainty due to the weighing process. The scatter in the data for different measurements represents the repeatability of the measurements. The level of the repeatability is probably caused by unstable masses due to adsorption of dirt, water etc. on the weight surfaces. The lines represent the average mass value and the 1σ confidence intervals from a least squares fit to a model describing drift and random variation. It is clear that no significant changes in mass values and calibration uncertainties are detected after the new software system is implemented.

We have also re-analyzed the raw data from an old calibration task made with the software previously in use, which was based on the analysis described in section 2 and 3. These raw data consist of weighing results (ΔI), values for environmental parameters and information on weights such as volumes, heights etc. The format of the raw data is adjusted to the format used for the new

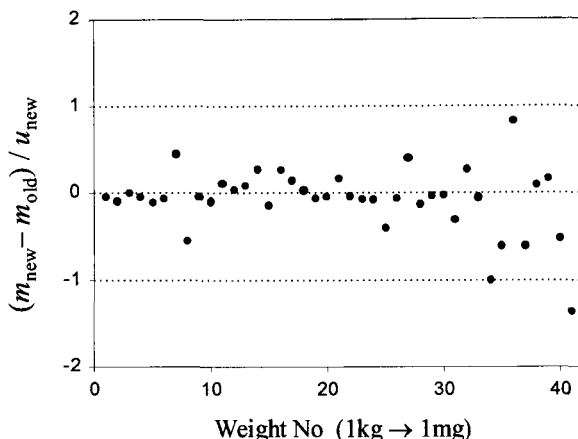


Fig. 3. Comparison of the calculated masses using the old and the new software/method for analysis. Both analyses use the same set of raw data. The mass differences are normalized to the calculated uncertainty.

software, and the data are reanalyzed using the new software with the method described in section 4. The results from the old analysis (i.e. original analysis with old software) and the new analysis are shown in Fig. 3. It is seen that the two software products using two different methods for the analysis agree within the calculated uncertainty. The increased scatter at larger weight numbers can be explained by insufficient resolution in the recorded m_{old} values at small mass values. Besides looking at the mass values, we have also investigated the uncertainties obtained with the old and new software.

The final test is based on the generation of an artificial test data set. Although the calculation of mass values from the weighing results is quite difficult in general, the reverse calculation of weighing results from known masses is in fact simple. Thus, we assume that the masses of all the weights are known for a simple design with 9 weights. From the assumed environmental values and the known weight volumes, heights etc. we calculate the corresponding weighing results (ΔI) using Eq. (2) and (3). These calculated weighing results will correspond to a “perfect fit”, and the subsequent analysis is therefore independent of the uncertainty on the weighing results. In Fig. 4 we show the comparison between the assumed input masses and the output masses from the analysis, and the agreement is quite good. The relative difference between input and output masses is at the 10^{-12} level simply due to the number of digits (12) in the output of the analysis program. For comparison, the uncertainty in the mass value of the national kilogram prototype is $\approx 3 \times 10^{-9}$.

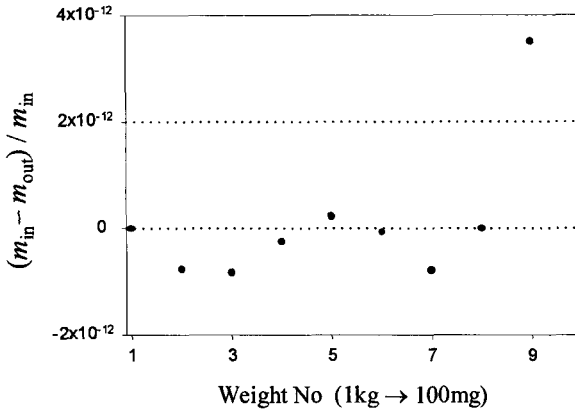


Fig. 4. Comparison of input masses used for generating the test data set and the output masses from the analysis of the test data.

7. Conclusion

In conclusion, we have described the principles of mass calibration and introduced the general method of least squares, which has been implemented in the new mass calibration software. The validation of the software has been discussed and results have been presented.

References

1. J.M. Cameron, M.C. Croarkin, R.C. Raybold, *Designs for the calibration of Standards of Mass*, NBS Technical Note 952 (1977)
2. W. Bich, *Metrologia* **40**, 306 (2003).
3. R. S. Davis, *Metrologia* **29**, 67 (1992).
4. P. Giacomo, *Metrologia* **18**, 33 (1982).
5. L. Nielsen, *Metrologia* **35**, 115 (1998). Erratum, *Metrologia* **37**, 183 (2000).
6. L. Nielsen, "Evaluation of measurements by the method of least squares", in *Algorithms for Approximation IV*, J. Levesley, I.J. Anderson, J.C. Mason (Eds), University of Huddersfield, 2002, 170-186.
7. OIML International Recommendation No. 33: *Conventional value for the result of weighing in air*, (1989).
8. Nordtest report (www.nordtest.org), NT Tech Report 535: *Development of test method for validation of software for calibration and testing*, (2003).

THE *GUM TREE* DESIGN PATTERN FOR UNCERTAINTY SOFTWARE

B. D. HALL

*Measurement Standards Laboratory of New Zealand,
Industrial Research Ltd., PO Box 31-310,
Lower Hutt, New Zealand
E-mail: b.hall@irl.cri.nz*

The paper describes a simple approach to software design in which the ‘Law of propagation of uncertainty’ is used to obtain measurement results that include a statement of uncertainty, as described in the *Guide to the Expression of Uncertainty in Measurement* (ISO, Geneva, 1995). The technique can be used directly for measurement uncertainty calculations, but is of particular interest when applied to the design of instrumentation systems. It supports modularity and extensibility, which are key requirements of modern instrumentation, without imposing an additional performance burden. The technique automates the evaluation and propagation of components of uncertainty in an arbitrary network of modular measurement components.

1. Introduction

This paper describes a design technique for software that applies the ‘Law of propagation of uncertainty’, as recommended in the *Guide to the Expression of Uncertainty in Measurement* (henceforth, the *Guide*)¹. It automates the task of evaluating and propagating components of uncertainty and produces software that is both modular and extensible, both of which are requirements for modern instrumentation.

The technique can be applied to instrument and smart-sensor firmware, or to higher level software coordinating measurement systems for specific measurements. It is both efficient and simple to implement. Its most striking feature is that measurement equations are differentiated automatically, allowing uncertainty to be propagated with only a direct statement of measurement equations.

The next section shows that the *Guide*’s formulation of the uncertainty propagation law can be easily applied to modular measurement systems. By decomposing the evaluation of a measurement value and the associated

uncertainty into an arbitrary set of intermediate calculations, we show that each step can be considered as a modular component of a system. Section 3 uses this formulation to identify the basic requirements of modular software components capable of encapsulating the calculations.

2. Calculation of value and uncertainty in modular systems

2.1. *The measurement function*

According to the *Guide*, a measurement may be modelled by a single ‘measurement function’, f , interpreted

“ ... as that function which contains every quantity, including all corrections and correction factors, that can contribute a significant component of uncertainty to the measurement result”¹.

The inputs to this function are uncertain quantities so a measurement result is estimated as

$$x_m = f(x_1, x_2, \dots, x_l), \quad (1)$$

where f is evaluated at the estimated mean values of the inputs, x_1, x_2, \dots, x_l .

For example, electrical power, P , dissipated in a resistor may be estimated by measuring the potential difference across the resistor terminals, V , and the resistance, R . The measurement function for this is

$$P = \frac{V^2}{R}, \quad (2)$$

where V and R are the two inputs.

2.2. *Decomposing the measurement function*

In the *Guide*’s approach, the function f represents the entire measurement procedure. In complex measurement systems it will be difficult to state its function explicitly. However, f can be conveniently decomposed into a set of simpler functions of the form

$$x_j = f_j(\Lambda_j), \quad (3)$$

where ‘ j ’ identifies the decomposition step, x_j is the intermediate value at the step and Λ_j is the set of direct inputs to the function f_j .^a These

^aThe labels, j , are assigned so that $j > k$, where k is the subscript of any member of the set Λ_j .

functions can be used to evaluate the measurement result by using a simple iterative algorithm. If there are $m - l$ decomposition steps and l direct inputs to f (called ‘system inputs’), then

$$x_j = f_j(\Lambda_j), \text{ for } j = l + 1, \dots, m \quad (4)$$

The decomposition steps of a measurement function are referred to from now on as ‘modules’ because we will show that both value and uncertainty calculations can actually be encapsulated at each step and thought of as a discrete component of a measurement. The algorithm above can be interpreted as a distributed calculation over the modules comprising a system.

Referring to the simple power measurement example, a possible decomposition is:

$$x_1 = V \quad (5)$$

$$x_2 = R \quad (6)$$

$$x_3 = x_1^2 \quad (7)$$

$$P = x_4 = x_3/x_2. \quad (8)$$

There are two system inputs, x_1 and x_2 , so $l = 2$, and two more modules that decompose equation (2), so $m = 4$.

2.3. *Standard uncertainty and components of uncertainty*

The term ‘standard uncertainty’, denoted $u(x_i)$, is an estimate of the standard deviation of the random variable associated with the quantity x_i . The standard uncertainty in the measurement result, $u(x_m)$, often referred to as the ‘combined standard uncertainty’, depends on the input uncertainties $u(x_1), u(x_2), \dots, u(x_l)$, their possible correlations and on the function f .

The calculation of $u(x_m)$ is simplified by defining components of uncertainty, which are related to the sensitivity of the measurement function to its inputs. The ‘component of uncertainty in x_m due to uncertainty in an input quantity x_i ’ is defined as^b

$$u_i(x_m) \equiv \frac{\partial f}{\partial x_i} u(x_i), \quad (9)$$

where the partial derivative is evaluated at the values of x_1, x_2, \dots, x_l .

^bThe *Guide* defines a component of uncertainty as the *modulus* of $u_i(x_m)$ here.

The combined standard uncertainty can be evaluated from the $u_i(x_m)$ and the associated correlation coefficients

$$u(x_m) = \left[\sum_{i=1}^l \sum_{j=1}^l u_i(x_m) r(x_i, x_j) u_j(x_m) \right]^{1/2}, \quad (10)$$

where the sums are over the l system inputs and $r(x_i, x_j)$ is the correlation coefficient between inputs x_i and x_j .

In terms of the power example, the components of uncertainty are

$$u_V(P) = u_1(x_4) = 2 \frac{V}{R} u(V) \quad (11)$$

$$u_R(P) = u_2(x_4) = -\frac{V^2}{R^2} u(R), \quad (12)$$

and if the measurements of V and R are uncorrelated, the combined standard uncertainty in the power measurement is

$$u(P) = \frac{V}{R} \left[4u^2(V) + \frac{V^2}{R^2} u^2(R) \right]^{1/2}. \quad (13)$$

2.4. Decomposition of the uncertainty components

The key point made in this paper is that evaluation of equation (9) can be decomposed into steps to obtain an algorithm of the same structure as (4). This makes it possible to encapsulate both the calculation of components of uncertainty, and the calculation of value, within modules. Each module need only evaluate the components of uncertainty of its own and the results will propagate through the system correctly.

Equation (9) can be decomposed by applying the chain rule for partial differentiation to the decomposition functions in algorithm (3). A component of uncertainty, for any given i , can then be calculated iteratively:

$$u_i(x_j) = \sum_{x_k \in \Lambda_j} \frac{\partial f_j}{\partial x_k} u_i(x_k), \text{ for } j = l+1, \dots, m. \quad (14)$$

This algorithm propagates the sensitivity of each module function, f_j , to a module input x_k , weighted by the component of uncertainty $u_i(x_k)$ in that input's value due to uncertainty in the system input x_i .^c

This algorithm, together with equation (4), show that the uncertainty component and value calculations of a step can be encapsulated in one

^cEquation (14) implies that $u_i(x_i) \equiv u(x_i)$. This is compatible with (9), when x_m is replaced by x_j and f by f_j .

module. The algorithms effectively distribute both value and uncertainty calculations over a network of interconnected modules.

In the power example, referring to equations (5)–(8), the two uncertainty functions ($j = 3$ and $j = 4$), for arbitrary i , are

$$u_i(x_3) = \frac{\partial x_3}{\partial x_1} u_i(x_1) = 2x_1 u_i(x_1) \quad (15)$$

$$\begin{aligned} u_i(x_4) &= \frac{\partial x_4}{\partial x_2} u_i(x_2) + \frac{\partial x_4}{\partial x_3} u_i(x_3) \\ &= -\frac{x_3}{x_2^2} u_i(x_2) + \frac{1}{x_2} u_i(x_3) . \end{aligned} \quad (16)$$

In these equations, together with equations (7) and (8), a module need only obtain information from its immediate inputs (the arguments to f_j). So the $j = 3$ module (equations (7) and (15)) only refers to x_1 (the voltage V) and the $j = 4$ module (equations (8) and (16)) only refers to x_3 (the voltage squared, V^2) and x_2 (the resistance, R).

3. A design pattern for uncertainty software

A design pattern describes a generic solution to a broad class of problem. Patterns are widely used in software development and are an effective way for software developers to communicate and share successful techniques. They provide flexibility (e.g. allowing implementation in various programming languages) while concisely describing a solution strategy. The *GUM Tree* design pattern, shows how to incorporate, into the design of measurement systems, the mathematics underlying uncertainty propagation.

The *GUM Tree* is closely related to the more general and well documented ‘Interpreter’ pattern². There are two important software classes in the *GUM Tree*: a `Module` interface class and a `Context` class. The former is a base class for system modules. It constrains a module’s implementation class to meet the basic requirements necessary for uncertainty propagation. The `Context` class encapsulates calculations that must be handled outside of modules and data associated with the overall system.

3.1. The requirements of a module interface

To evaluate algorithms (4) and (14) the j^{th} module must provide two pieces of information:

- an output value, x_j ;

- a component of uncertainty in the (module) output due to uncertainty in the i^{th} system input, $u_i(x_j)$.

Every module's software interface should include functions that return these numbers.^d In C++ pseudo code, such an interface can be declared as a pure abstract base class^e

```
class ModuleInterface {
public:
    virtual double value() = 0;
    virtual double uComponent(ModuleInterface& i) = 0;
};
```

3.2. *The Context*

A separate class is required, in addition to the library of module classes, to encapsulate global calculations and manage global data. For example, the calculation of combined standard uncertainty, equation (10), has to be performed on the system modules and needs access to correlation coefficients. An entity called the 'Context' can be used to encapsulate this type of information.^f The Context is also useful for handling language and implementation specific details and can extend native language features if necessary. For instance, it can provide support for memory management and expression parsing with abstract data types.

3.3. *Applying the pattern*

Implementation of the GUM Tree pattern requires that a set module classes be defined, each of which implements the module interface in a different way. Module objects can be created from this class library and linked together to form a system. When in operation, a Context will interrogate the network of modules to determine measured values and their uncertainties.

Equations (4) require only that a module be characterised by a simple function, f_j , which may have an arbitrary number of arguments. Hence

^dOther module interface functions will generally be useful. For instance, an interface can include a function that returns an iterator for the set of system inputs on which the module depends directly or indirectly, this helps sum over i and j in equation (10).

^eIn practice, a distinct interface class, derived from `ModuleInterface`, can be useful to distinguish between system inputs and module inputs. This could be used as the argument type in the second function declaration.

^fAnother calculation that the Context would typically handle is the evaluation of effective degrees-of-freedom¹.

modules may be distinct pieces of instrumentation, connected to the system by some kind of physical communications interface, or they may be abstract software entities in computer memory.

An example of the latter is a class library for uncertainty calculations⁶. This effectively introduces a new abstract data type for measurement uncertainty calculations – an entity that has attributes for both value and components of uncertainty.⁸ Module classes can implement simple maths functions and arithmetic operations, allowing arbitrary measurement functions to be written directly in software. The explicit derivation of sensitivity coefficients is then unnecessary: a measurement function will be parsed automatically into a tree of linked module objects, which can be interrogated to obtain values and uncertainties.

As an indication of how simple such a class library is, consider the definition of a module class for division. The function definitions were given in equations (8) and 16. In addition, the class will need two ‘references’ (x_2 and x_3), which link to a module’s inputs. Member functions use these references to obtain information (i.e., an output value or an uncertainty component). In C++ pseudo-code, the class definition for division looks like⁶

```
class Division : public ModuleInterface {
protected:
    ModuleInterface& x2;
    ModuleInterface& x3;

public:
    Division(ModuleInterface& l, ModuleInterface& r) x3(l), x2(r) {}

    double value() { return x3.value() / x2.value(); }

    double uComponent(ModuleInterface& i) {
        double v = x3.value() / x2.value();
        return
            2 * v * x3.uComponent(i) -
            v / x2.value() * x2.uComponent(i);
    }
};
```

⁸Much like complex numbers, and mathematical support for them, can be added to some programming languages.

The member function `value()` corresponds to equation (8) and the member function `uComponent(ModuleInterface& i)` to equation (16) with `i` a reference to any system input module. The class constructor, `Division(ModuleInterface& l, ModuleInterface& r)`, establishes the required links to the module inputs. This type of implementation of the GUM Tree is closely related to the 'reverse' form of automatic differentiation⁵.

The extensibility of GUM Tree designs relies on the common `Module` interface. For example, the input links in the pseudo code above gave no hint as to the extent of the module network connected to each input. A link could point to a simple input or to a complicated module network. This generality allows exchange and extension to occur dynamically in a system.

For instance, equation (6) attributed a system input value of R to the measurement equation. However, as a modification to the system, a temperature dependent resistance

$$R = x_2 = R_0[1 - \alpha(T - T_0)] \quad (17)$$

could be associated with that step in the calculation. The network of modules representing the new resistance calculation are then connected at this point (i.e., a decomposition tree of modules representing equation (17)). This change does not require any changes to code in other modules.^h

4. Discussion and Conclusions

The GUM Tree design pattern is a simple technique that represents a significant improvement over current practice and should ultimately result in more reliable and robust measurement and control systems throughout society. It could be deployed in most modern instrument designs where it would allow systems to report measured values together with uncertainties.

Current trends in the test and measurement industry are towards intelligent and flexible system components with industry-standard interfaces to facilitate exchange, re-configuration and long-term maintenance requirements^{3,4}. The modularity and extensibility of the GUM Tree pattern is fully compatible with this trend and has the potential to substantially enhance the expected benefits. The GUM Tree design enables systems

^hThe calculation of combined standard uncertainty will naturally be different and will include the additional uncertainty terms. However, this calculation can easily be made generic and parameterised on information available from the module interface.

to dynamically report uncertainty as operational parameters change, which greatly enhances the integrity of a system's measurement results. One interesting possibility is to reduce the cost-of-ownership burden associated with revalidating test and measurement systems when components are changed.

The information required for the uncertainty function of the GUM Tree module interface would be known during the design of an instrument or sensor, so there will be little difficulty in providing software to report it. On the other hand, legacy instruments could be endowed with a GUM Tree compatible interface, by introducing an additional software layer similar to the 'Role control module' approach described in ³.

The new software data type, representing measurement results by encapsulating value and uncertainty information, is particularly interesting. It could be used directly as a convenient tool for data processing (an alternative to 'uncertainty calculator' applications). However it is potentially more powerful in the internal software of measuring instruments. An instrument's calculations would then be following the *Guide's* recommendations and handling uncertainty at a very low level. This would not compromise proprietary algorithms, but would give external access to the sensitivity coefficients required for uncertainty calculations.

In conclusion, the GUM Tree design pattern is an elegant approach to instrument software design that allows the 'Law of propagation of uncertainty' to obtain measurement results that include a statement of uncertainty, as described in the *Guide*. The technique is applicable to measurement systems of many kinds, large and small. It supports modularity and extensibility, which are key requirements of modern instrumentation systems, and does not impose significant performance requirements. There is increasing pressure on test and measurement manufacturers to provide support for uncertainty that follows the recommendations of the *Guide*. The GUM Tree design pattern offers a solution to this problem.

Acknowledgements

The author is grateful to R. Willink, for many helpful discussions regarding the GUM Tree and measurement uncertainty in general, and to T. Armstrong for careful reading of this manuscript. Aspects of the GUM Tree design technique have been filed in a PCT application: *Uncertainty propagation system and method* (PCT/NZ02/00107, 6 June 2002).

References

1. ISO, *Guide to the expression of uncertainty in measurement*, International Organisation for Standardization, Geneva, 2nd edition, 1995.
2. E. Gamma, R. Helm, R. Johnson and J. Vlissides, *Design Patterns: Elements of reusable object-oriented software*, Addison-Wesley, Boston, 1995.
3. J. Mueller and R. Oblad "Architecture drives test system standards" *IEEE Spectrum* **September** (2000) 68.
4. K. L. Lee and R. D. Schneeman, "Internet-based distributed measurement and control applications," *IEEE Instrum. Meas. Mag.*, pp. 23–27, June 1999; T. R. Licht, "The IEEE 1451.4 proposed standard," *IEEE Instrum. Meas. Mag.*, pp. 12–18, March 2001.
5. L. B. Rall and G. F. Corliss, "An introduction to automatic differentiation," in *Computational Differentiation: Techniques Applications, and Tools*, M. Berz, C. H. Bischof, G. F. Corliss, and A. Griewankpp, Eds., pp. 1–17. SIAM, Philadelphia, September 1996.
6. A number of documents that describe the GUM Tree in more detail and give examples of its implementation are available for download at <http://www.irl.cri.nz/msl/mst>. The available material includes software libraries and examples.

STATISTICAL HYPOTHESES TESTING FOR PHASE TRANSITION IDENTIFICATION IN CRYOGENIC THERMOMETRY*

D. ICHIM, I. PERONI, F. SPARASCI

*Istituto di Metrologia "G. Colonnetti" - CNR,
Strada delle Cacce, 73
10135 Torino, Italy*

E-mail: d.ichim@imgc.cnr.it, i.peroni@imgc.cnr.it, f.sparasci@imgc.cnr.it

The identification of the phase transition is required in thermometry in order to reproduce the International Temperature Scale of 1990 (ITS-90) defining fixed points.

This paper proposes the use of statistical hypotheses testing methodologies for phase transition identification by simply monitoring the temperature behaviour. Only the pure structural change model is taken into consideration, that is the model in which the components of the parameter vector are allowed to change all together.

A statistical test for a single known change point is briefly presented in the paper. The procedure is extended to the more interesting application of the detection of a single unknown change point. A sliding window algorithm is proposed for the on-line detection of a possible change point.

The method has been applied to identify the triple points of the four gases realizing the cryogenic range of the ITS-90. The pure structural change model has proved to be an innovative tool for a reproducible and reliable phase transition identification.

1. Introduction

It is generally recognized that a physical quantity experiences changes as it evolves over time, especially over a long period. In a parametric regression framework, structural change may be defined as a change in one or more parameters of the model in question. This can also be seen as a parameter instability problem. Deeper knowledge on the parameter behaviour gives more information on whether a particular model is correct or not. Poor forecasts may result if the model does not account for structural change.

*Work partially funded under EU SofTools_MetroNet Contract N. G6RT-CT-2001-05061

There are a lot of examples in engineering, physics, as well as in metrology, where the identification of a defined phenomenon can be done by monitoring the parameter changes. It is about the processes that change from one regime to another, in a more or less gradual fashion, as in the case of the first order phase transition detection. In this particular case several physical quantities, such as entropy and volume show abrupt changes. On the other hand, phase transition occurrences correspond to well defined temperature states that can be used as fixed points as recommended in the International Temperature Scale of 1990 (ITS-90) [1]. Therefore, the identification of a phase transition is one of the fundamental aspects of thermometry.

The procedures currently used to realize a fixed point generally require the judgement of a trained operator, to decide when the phase transition is reached. Because phase transition measurements are extremely time consuming and require highly reproducible conditions, it is more convenient to carry out these measurements automatically. During a constant power approach to the transition, the temperature dependence on time abruptly changes its shape, as soon as the phase transition occurs. A temperature plateau happens and the mathematical model parameters change in a very definite way. To detect this change the statistical test presented below has been employed.

1.1. Assessed problem

Suppose we deal with a sequence of time ordered data. That is, at the moment n , $y_n \in \mathbb{R}$ and $\mathbf{x}_n = (x_{n,1}, \dots, x_{n,p})' \in \mathbb{R}^p$ are acquired, $n = 1, \dots, N$, where N is the sample size. Considering that Y is the response variable, denote by $\mathbf{Y} = (y_1, \dots, y_N)' \in \mathbb{R}^N$ the column vector of its observed values. Consider that X is the explanatory variable and denote by $\mathbf{X} = \begin{pmatrix} \mathbf{x}_1' & \mathbf{x}_2' & \dots & \mathbf{x}_N' \end{pmatrix} \in \mathbb{R}^{N,p}$ the matrix of its observed values.

Assume that the studied phenomenon can be described by a linear regression model:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\beta} &= (\beta_1, \beta_2, \dots, \beta_p)' \in \mathbb{R}^p \\ \boldsymbol{\varepsilon} &= (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)' \in \mathbb{R}^N, i.i.d. \sim \mathcal{N}(0, \sigma) \end{aligned} \quad (1)$$

The main goal of the regression analysis is to find the best estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$

with respect to a chosen norm. That is, find $\hat{\beta}$ such that:

$$\min_{\beta} \|Y - X\beta\| = \|Y - X\hat{\beta}\| \quad (2)$$

In this work the L_2 norm was used, that is:

$$\begin{aligned} \min_{\beta} \|Y - X\beta\| &= \min_{\beta} (Y - X\beta)' (Y - X\beta) \\ &= (Y - X\hat{\beta})' (Y - X\hat{\beta}) \end{aligned} \quad (3)$$

The problem assessed in this paper is the following: shall β be inferred using the entire sample of data or would it be more appropriate to split the data in two subsets and estimate the different subset parameters β_1, β_2, \dots ? For example, looking at the simulated data in the figure 1 it is quite obvious that more correct information would be obtained if the data were split in two subsets around the point $x = 32$.

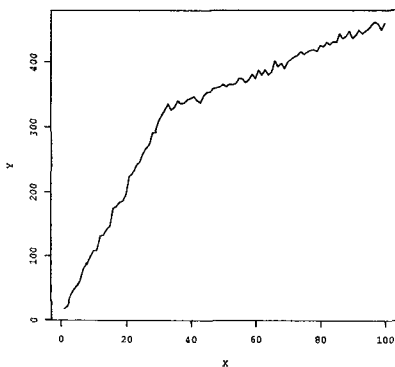


Figure 1. Simulation of a process showing a change of the model parameters around the point $x = 32$.

2. Structural Change Model

Consider the following linear regression model with a change point at N_0 :

$$y_n = \begin{cases} \mathbf{x}_n' \beta_1 + \varepsilon_n, & n = 1, \dots, N_0 \\ \mathbf{x}_n' \beta_2 + \varepsilon_n, & n = N_0 + 1, \dots, N \end{cases} \quad (4)$$

where ε_n are i.i.d, $\varepsilon_n \sim \mathcal{N}(0, \sigma)$, $\beta_1 = (\beta_{11}, \dots, \beta_{1p})' \in \mathbb{R}^p$ and $\beta_2 = (\beta_{21}, \dots, \beta_{2p})' \in \mathbb{R}^p$ for some known or unknown change point

$N_0 \in \{2, \dots, N-1\}$. Assume that the matrices $\sum_{n=1}^{N_0} \mathbf{x}_n \mathbf{x}_n'$ and $\sum_{n=N_0+1}^N \mathbf{x}_n \mathbf{x}_n'$ are both of full rank p .

In this paper we take into consideration only the case of a pure structural change model, where all the components of the parameter β are allowed to change and checked for variation all together.

There are four statistical items to be discussed when considering a structural change model. The first point is to assess whether a change occurred or not. That is, by statistical means, decide whether there is evidence to use a model like (4) to perform the required inferences, rather than the classical model (1) for the same purposes. In case we decide for a change occurrence, the second phase is to estimate when the change happened in the studied data set, for an unknown change point. The rest of the statistical procedure is well known: thirdly, estimate the parameters and, fourthly, forecast future values. In this paper we discuss only the first item: assessing the occurrence of a change in the framework of a single possible change point.

2.1. Parametric hypotheses testing

In order to decide for change in the analyzed parameters, a statistical test must be performed. The main ingredients of a statistical test are the null hypothesis H_0 and the alternative hypothesis H_1 . H_0 is the statement we trust unless the sample observations provide evidence favourable to the alternative statement H_1 . Testing H_0 against H_1 leads in fact to only two possible conclusions: a) "reject H_0 " and b) "do not reject H_0 ", without any inference about H_1 .

Another choice the operator has to make is the test significance level α , that is the probability of rejecting the null hypothesis H_0 when it is true. This probability α is also called probability of 1st type error, see [2].

A statistical test is performed by means of a test statistic T . T is a statistic whose distribution is known when the parameter β is known. T is generally determined on the basis of the parameters to be tested and of their distribution [2].

With the above mentioned items, a critical region CR is defined. CR is an interval containing possible values of T such that $\alpha = P(T \in CR | H_0 \text{ is true})$. A best critical region CR^* is to be chosen, that is a CR with the following additional property:

$$\gamma = P(T \notin CR^* | H_1 \text{ is true}) = \min_{CR} P(T \notin CR | H_1 \text{ is true}) \quad (5)$$

γ is also called probability of 2^{nd} type error.

The decision provided by a statistical test is then defined by:

$$\begin{aligned} T \in CR^* &\implies \text{reject } H_0 \\ T \notin CR^* &\implies \text{do not reject } H_0 \end{aligned} \quad (6)$$

2.2. Structural change hypotheses testing: the known change point case

In case of a known change point $N_0 \in \{2, \dots, N-1\}$, the test to be performed in order to decide whether a change occurred or not is the following:

$$\begin{aligned} H_0 : \quad &\beta_1 = \beta_2 \\ \text{against} \quad & \\ H_1 : \quad &\beta_1 \neq \beta_2 \end{aligned} \quad (7)$$

when the notations used in (4) are still kept. It is here that the assumption of a pure structural change model appears.

Using the ordinary least squares (OLS) method as an estimation method, denote by RSS the sum of squared residuals obtained by solving the model (1) and, respectively, denote by RSS_{12} the sum of the two sums of squared residuals obtained by performing the two distinct regressions corresponding to the model (4). That is,

$$\begin{aligned} RSS &= \mathbf{Y}' \left[I - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] \mathbf{Y} \\ RSS_{12} &= \mathbf{Y}'_1 \left[I - \mathbf{X}_1(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1 \right] \mathbf{Y}_1 + \mathbf{Y}'_2 \left[I - \mathbf{X}_2(\mathbf{X}'_2\mathbf{X}_2)^{-1}\mathbf{X}'_2 \right] \mathbf{Y}_2 \end{aligned} \quad (8)$$

where

$$\mathbf{Y}_1 = \begin{bmatrix} y_1 \\ \dots \\ y_{N_0} \end{bmatrix}; \mathbf{Y}_2 = \begin{bmatrix} y_{N_0+1} \\ \dots \\ y_N \end{bmatrix}; \mathbf{X}_1 = \begin{bmatrix} \mathbf{X}_{1,.} \\ \dots \\ \mathbf{X}_{N_0,.} \end{bmatrix}; \mathbf{X}_2 = \begin{bmatrix} \mathbf{X}_{N_0+1,.} \\ \dots \\ \mathbf{X}_{N,.} \end{bmatrix} \quad (9)$$

and $A_{i,.}$ denotes the i -th row of a matrix A .

G. Chow [3] proved that the test statistic for (7) is given by:

$$T_{Chow} = \frac{(RSS - RSS_{12})/p}{RSS_{12}/(N - 2p)} \quad (10)$$

which has an exact F distribution with p and $N - 2p$ degrees of freedom respectively.

Applying the Chow test to the data presented in the figure 1, the test statistic T takes the value 6511.7, while the 99% critical value (defining

the critical region) is 4.7. As expected, the decision to be taken is that a change in the parameter values must be considered.

2.3. Structural change hypotheses testing: the unknown change point case

Consider now only the simplest case of a single unknown change point:

$$y_n = \begin{cases} \mathbf{x}'_n \boldsymbol{\beta} + \varepsilon_n, & n = 1, \dots, N_0 \\ \mathbf{x}'_n \boldsymbol{\beta} + \mathbf{x}'_n \boldsymbol{\beta}_1 + \varepsilon_n, & n = N_0 + 1, \dots, N \end{cases} \quad (11)$$

where $N_0 \in \mathcal{N}$. $\mathcal{N} \subset \{2, \dots, N-1\}$ will be called "inspection interval" in this paper. Later we'll see why it is necessary to introduce it.

Then the statistical test to be performed is expressed as:

$$\begin{aligned} H_0 : & \quad \boldsymbol{\beta}_1 = 0 \\ \text{against} & \\ H_1 : & \quad \boldsymbol{\beta}_1 \neq 0 \text{ for some } N_0. \end{aligned} \quad (12)$$

Testing for structural change with unknown change point does not fit into the regular testing framework because the parameter N_0 only appears under the alternative hypothesis and not under the null one H_0 . In consequence, the Wald, the Lagrange multiplier and the likelihood-ratio like tests do not possess their standard large sample asymptotic distribution and hence cannot be used.

Denoting by:

$$\begin{aligned} \mathbf{X}_{2n} &= \begin{pmatrix} \mathbf{0}' : \mathbf{0}' : \dots : \mathbf{0}' : \mathbf{x}'_{n+1} : \dots : \mathbf{x}'_N \end{pmatrix} \in \mathbb{R}^{N,p} \\ \mathbf{X}_0 &= \begin{pmatrix} \mathbf{0}' : \dots : \mathbf{0}' : \mathbf{x}'_{N_0+1} : \dots : \mathbf{x}'_N \end{pmatrix} \in \mathbb{R}^{N,p} \end{aligned} \quad (13)$$

the model (11) can be rewritten as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}_0\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon} \quad (14)$$

Let $RSS(n)$ be the sum of squared residuals and let $(\hat{\boldsymbol{\beta}}_n, \hat{\boldsymbol{\beta}}_{1n})$ be the OLS estimator of $(\boldsymbol{\beta}, \boldsymbol{\beta}_1)$ by regressing \mathbf{Y} on \mathbf{X} and \mathbf{X}_{2n} . In the case of unknown variance σ^2 , it can be estimated by $\hat{\sigma}^2(n) = RSS(n)/(N-2p)$. The test statistic for the unknown single change point model is then constructed simply by generalizing the known change point case by defining first $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ and $W(n/N) = \left(\hat{\boldsymbol{\beta}}'_{1n} \left(\mathbf{X}'_{2n} \mathbf{M} \mathbf{X}_{2n} \right) \hat{\boldsymbol{\beta}}_{1n} \right) / \hat{\sigma}^2(n)$.

Then:

$$T = \sup_{n \in \mathcal{N}} W\left(\frac{n}{N}\right) \quad (15)$$

D. Andrews in [4] founded the asymptotic distribution of this test statistic. Note that in the case of the known change point model the exact distribution was found, while in this case only the asymptotic one can be obtained. Let $B_p(\cdot)$ be a p -vector of independent Brownian motions on $[0,1]$ restricted to the inspection interval \mathcal{N} . In [4], calling by $Q_p(\pi) = \left((B_p(\pi) - \pi B_p(1))' (B_p(\pi) - \pi B_p(1)) \right) / (\pi(1 - \pi))$ a standardized tied-down Bessel process, Andrews proved that the statistic $W(\cdot)$ weakly converges to $Q_p(\cdot)$, while $T = \sup_{n \in \mathcal{N}} W(\cdot)$ converges in distribution to $\sup_{\mathcal{N}} Q_p(\cdot)$. He also proved that the test statistic does not converge on the entire interval. This seems to be reasonable as a change at the extremes of the interval would mean in fact “no change at all”. This is the reason why the inspection interval must be kept away from the ends of the data interval. The critical values for the statistic $\sup_{\mathcal{N}} Q_p(\cdot)$, with respect to various inspection intervals \mathcal{N} , are obtained in the literature by Monte Carlo simulations; see [4] for example. An estimation \hat{N}_0 of the change point N_0 is to be defined by the point for which the supremum (15) is reached. For any confidence level, confidence intervals can be constructed as in [5], for example.

3. Application to phase transition identification in cryogenic thermometry

We applied the statistical method presented in the previous section to identify the triple points of the gases realizing the cryogenic range of ITS-90. The phase transition identification allows us to recognize a physical state corresponding to a well defined temperature that we want to reproduce. When the phase transition is approached, supplying a constant power to the system, the transition occurrence is characterized by a sudden change in temperature versus time behaviour, which corresponds to discontinuities in some physical parameters. As soon as the transition happens, the melting process starts and a temperature plateau occurs. The rapid identification of the change gives information on the entire melting process.

In fact, by identifying the extremes of the melting, the total heat of fusion and the transition temperature at different melted gas fractions are determined. These parameter values are more accurately obtained when a “real-time” identification of the starting of the fusion process is performed.

Using the statistical test previously presented, together with a simple temperature monitoring, we detected the phase transition by means of a completely automatic procedure. The identification was performed on-line

with a negligible delay with respect to the plateau duration. We actually applied this statistical test to the triple points of H_2 , Ne, O_2 and Ar, covering the cryogenic temperature range.

In cryogenic metrology other techniques, mainly based on discontinuity detection, are used for phase transition identification. All these methods generally perform approximations and computations for each data point acquired.

In our case only the temperature monitoring is required: there is no need to know specific characteristics of the phase transition in study. In contrast with other techniques, no measurements of physical quantities, except temperature, must be performed to investigate the occurrence of the transition.

In particular, phase transitions happening at unknown temperatures can be recognized. This means that the same technique, practically with the same parameters, can be applied for different phase transitions. Only the power supplied to the system must be scaled with temperatures.

3.1. *On-line change detection*

The change detection algorithms are mainly designed to be used off-line because they require the analysis of the entire data set. We used a sliding window procedure for the on-line detection of the phase transition. The

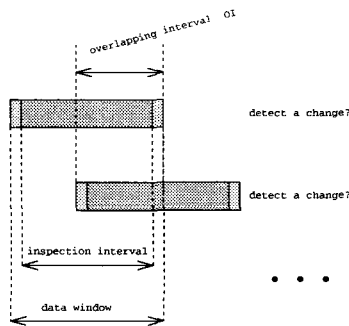


Figure 2. The sliding window procedure.

statistical analysis was implemented in Matlab® and inserted in the overall

LabVIEWTM data acquisition software. Anyway, it has to be noted that, being an algorithm, our statistical test can be written in any programming language. Moreover, being based on OLS, in a very extreme simplification, it needs only basic arithmetic operations in order to be implemented.

In figure 2 the sliding window scheme is presented. As specified in the previous sections, the test statistic (15) does not converge on the entire interval. In practice this means that a check for change corresponding to all points cannot be performed. In order to be able to check all the data for change, the data windows must be overlapped. Let the overlapping interval, OI , be the intersection of two consecutive data windows. Unfortunately the overlapping introduces a delay in the change detection as we can state the change (or not) only at the end of the data window we currently analyze. On the other hand, not checking all the data means to accept the risk of never detecting the change, which, at least in our case, is more dangerous than the previous situation. The parameters to be defined when using the proposed

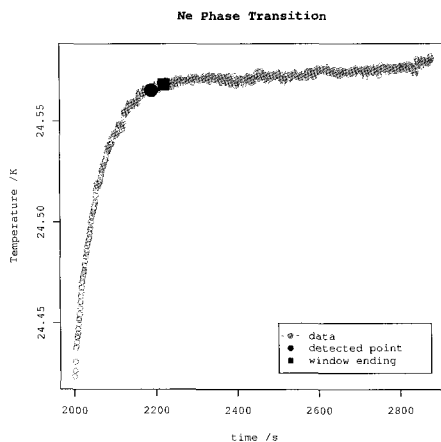


Figure 3. Example of a Neon triple point realization. The data, the detected change point and the end of the window containing the detected change point are indicated. Parameter values: $N = 100$, $\alpha = 0.05$, $\mathcal{N} = 70(\text{centered})$, $OI = 50$.

sliding window procedure are the statistical test level α , the window size N , the inspection interval \mathcal{N} and the overlapping interval OI . We varied these parameters in order to discover how sensitive our problem is to their values. Needless to say, the choice of the test significance level α can be made arbitrarily, also because the kind of change we are dealing with is so

abrupt that it should be detected at any significance level. There is a strong relationship among the other three parameters. The greater the window size, the longer the inspection interval can be, but the analysis tends to an off-line one. The greater the ratio N/N is, the smaller the overlapping interval can be, and hence the faster the detection. We observed that the detection is not at all sensitive to the variation of these parameters, in the sense that the change is always detected. The only sensitivity is related to the delay introduced by the parameters choice. In figure 3 an example of a phase transition detection of a Neon triple point realization is shown. Using an acquisition rate of 1 s, with $N = 100$, $N = 70$ (centered), $OI = 50$, $\alpha = 0.05$, the greatest delay we obtained in this application was about 1 minute. Further studies on these parameters should be performed in order to minimize the delay.

4. Conclusions

The statistical test presented proved to be a powerful tool for determining whether a pure structural change has taken place. Furthermore, it characterizes the detected change by providing confidence levels and confidence intervals. Particular features of the test presented in this paper are the low computational cost and its flexibility. The procedure discussed here is based on least squares computations, and hence is not very burdensome computationally. On the other hand, the flexibility is shown by the fact that we can apply the same test to different kinds of data: phase transition identifications, change in sensor calibration curves, study of ageing of materials, etc.

The analyzed data proved that the structural change point model is a reproducible and reliable method for the identification of the phase transition of the four gas triple points realizing the cryogenic range of ITS-90.

In future work we intend to compare the pure structural change model to a partial change model and to a gradual change model, all of them being possible candidates to our phase transition problem.

References

1. *Supplementary information for the International Temperature Scale of 1990*, BIPM, (1990)
2. E. L. Lehmann, *Testing Statistical Hypotheses*, (1959).
3. G. C. Chow, *Econometrica* **28**, 591 (1960).
4. D. W. K. Andrews, *Econometrica* **61**, 821 (1993).
5. J. Bai, *Review of Economics and Statistics* **79**, 551 (1997).

THE IMPACT OF ENTROPY OPTIMIZATION PRINCIPLES ON THE PROBABILITY ASSIGNMENT TO THE MEASUREMENT UNCERTAINTY

G. IUCULANO AND A. ZANOBINI

Department of Electronics Engineering, University of Florence,

Via di S. Marta, 3 – Florence – I 50139

E-mail: iuculano@ingfil.ing.unifi.it

G. PELLEGRINI

Department of Applied Mathematics, Engineering Faculty,

University of Florence, Via di S. Marta, 3 – Florence – I 50139

E-mail: gualtieri@dma.unifi.it

A measurement process has imperfections that give rise to uncertainty in each measurement result. Statistical tools give the assessment of uncertainties associated to the results only if all the relevant quantities involved in the process are interpreted or regarded as random variables. In other terms all the sources of uncertainty are characterized by probability distribution functions, the form of which is assumed to either be known from measurements or unknown and so conjectured. Entropy is an information measure associated with the probability distribution of any random variable, so that it plays an important role in the metrological activity.

In this paper the authors introduce two basic entropy optimization principles: the Jaynes's principle of maximum entropy and the Kulback's principle of minimum cross-entropy (minimum directed divergence) and discuss the methods to approach the optimal solution of those entropic forms in some specific measurements models.

1. Introduction

The statement of measurement uncertainty is not complete by merely stating point estimates, but usually is given by a coverage interval (or region) estimation associated with the results of a measurement, that characterize the dispersion of the values that could reasonably be attributed to the measurand, that is the quantity subjected to measurement [1].

This goal is achievable only if there are basic probability assignments to all the relevant quantities involved in the measurement process and if it is assumed that one has some empirical knowledge of their distributional characteristic.

To fix the form of a distribution a rather personalistic point of view, which takes into account the available nonstatistical knowledge could be introduced.

It is clearly preferable to have an objective probability assignment because different researchers (laboratories) would infer the same distribution from the data associated with the same measurand obtained in the same way and in the same circumstances.

The application of the Jaynes's principle of maximum entropy (pme) could satisfy this exigency. In fact the concept of entropy is closely tied to the concept of uncertainty embedded in probability distributions: entropy can be defined as a measure of probabilistic uncertainty.

In particular if we had to choose a probability distribution for a chance experiment (measurement) without any prior knowledge about that distribution, it would seem reasonable to pick the uniform distribution because we have no reason to choose any other (as stated by the Principle of Insufficient Reason) and because that distribution maximizes the "uncertainty" of the outcome. But what if we had some prior knowledge of the distribution?

Suppose for example that we know of some constraints that the moments of that distribution have to satisfy. It is here that Kullback's principle of minimum cross entropy, derived from Shannon's measure of uncertainty or entropy, plays its role.

The principle of minimum cross-entropy provides a general method of inference about an unknown probability density when there exists a prior estimate of the density and new information in the form of constraints on expected values. The principle states that, of all densities that satisfy the constraints, one should choose the one with the least cross-entropy.

It is important to emphasize that minimizing cross entropy is equivalent to maximizing entropy when the prior one is a uniform distribution.

In conclusion, a measurement process can be regarded as a controlled learning process in which various aspects on uncertainty analysis are investigated and the substantial amount of information, got with respect to the conditions prior to the result after the measurement process is performed, so that it can be connected to the entropy optimization principles, that are correct methods of inductive inference when no sufficient knowledge about the statistical distributions of the involved random variables is available before the measurement process is carried out except for the permitted ranges, the essential model relationships and some constraints, gained in past experience, valuable usually in terms of expectations of given functions or bounds on them.

In this paper, by introducing the Jaynes's principle of maximum entropy and the Kullback's principle of minimum cross-entropy (minimum directed divergence), the authors discuss the methods to approach the optimal solution of those entropic forms in some specific measurements models.

2. The Entropy Optimization Principles and the Measurement Models

Entropy is an information measure, in this paper we are mainly concerned with the entropy of a continuous distribution defined by [2]:

$$H(f) = - \int_{-\infty}^{\infty} f(m) \ln f(m) dm \quad (1)$$

being $f(m)$ the density distribution function of a random variable M and assuming $H(f) \equiv 0$ if $f = 0$.

The principle of *maximum entropy* (MAXENT) states that, given a prescribed set of values μ_r , $r = 0, 1, 2, \dots$, of all the unknown distributions $g(m)$ (may be infinite) that satisfy some constraints as:

$$\mu_r = \int_{-\infty}^{\infty} m^r g(m) dm \quad r = 0, 1, 2, \dots \quad (2)$$

the best estimate of $g(m)$ is the one with the largest entropy (1).

The principle of *minimum cross-entropy* (MINCENT) is a generalization that applies in cases when a prior distribution, say $p(m) \neq 0$, that estimates $g(m)$ is known in addition to the constraints (2).

The principle state that of all the distributions $g(m) \neq 0$ that satisfy the constraints, you should choose the one having the least cross entropy defined as:

$$H(g, p) = \int_{-\infty}^{\infty} g(m) \ln \frac{g(m)}{p(m)} dm \quad (3)$$

From the operative point of view the Lagrange's multipliers method could be applied to determine the best estimate of $g(m)$ in both cases.

2.1. Application of the MAXENT to a measurement model

Given a quoted measurement uncertainty represented by the interval $[a, b]$, let M be a random variable in the interval $[a, b]$ representing the measure, whose statistical parameters are: the expected value $E\{M\}$ and the variance $Var\{M\}$.

The following relation is satisfied $E\{M^2\} = E^2\{M\} + Var\{M\}$.

Let S be the family of probability distributions every member of which, say $f(m)$ the generic one, is consistent with the constraints:

$$\int_a^b f(m) dm = 1; \quad \int_a^b m f(m) dm = E\{M\}; \quad \int_a^b m^2 f(m) dm = Var\{M\} + E^2\{M\} \quad (4)$$

We have to determine in S that distribution whose entropy is greater than that of every other member of the family, using Lagrange's multipliers method.

Taking into account equations (1) and (4), the Lagrangian is written:

$$L = - \int_a^b f(m) \ln f(m) dm - \lambda_0 \left[\int_a^b f(m) dm - 1 \right] - \lambda_1 \left[\int_a^b m f(m) dm - E\{M\} \right] - \lambda_2 \left\{ \int_a^b m^2 f(m) dm - [Var\{M\} + E^2\{M\}] \right\}$$

being $\lambda_0, \lambda_1, \lambda_2$ the Lagrange's multipliers.

From the Eulero-Lagrange equation $-\ln f(m) - 1 - \lambda_0 - \lambda_1 m - \lambda_2 m^2 = 0$ we get

$$f(m) = e^{-(1+\lambda_0)} e^{-\lambda_1 m - \lambda_2 m^2} = e^{-(1+\lambda_0)} e^{\frac{\lambda_1^2 - \lambda_2 \left(m + \frac{\lambda_1}{2\lambda_2}\right)^2}{4\lambda_2}}; \lambda_2 > 0 \quad (5)$$

By introducing the following transformation on the Lagrange's multipliers:

$$A = e^{-(1+\lambda_0)} e^{\frac{\lambda_1^2}{4\lambda_2}} \sqrt{\frac{\pi}{\lambda_2}}; \quad \frac{\lambda_1}{2\lambda_2} = -\mu; \quad \lambda_2 = \frac{1}{2\sigma^2}; \quad A > 0, \quad \sigma > 0$$

we can write:

$$f(m) = A \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(m-\mu)^2}{2\sigma^2}} \quad (6)$$

where

$$A^{-1} = \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{(m-\mu)^2}{2\sigma^2}} dm = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = P\{N(0,1) \leq x\} \quad (7)$$

By introducing the reduced normal density: $Z(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2}} = Z(-x)$

we can write:

$$f(m) = \frac{1}{\sigma} \left[\Phi(\bar{b}) - \Phi(\bar{a}) \right]^{-1} Z(\bar{m}) \quad (8)$$

being $\bar{\xi} = \frac{\xi - \mu}{\sigma}$, $\xi = a, b, m$.

The unknown μ and σ are determined by considering the 2nd and 3rd constraints in (4), so we have:

$$\mu - \sigma \left[\Phi(\bar{b}) - \Phi(\bar{a}) \right]^{-1} \left[Z(\bar{b}) - Z(\bar{a}) \right] = E\{M\} \quad (9)$$

$$\left\{ 1 + \frac{\left[\bar{a} \right] Z(\bar{a}) - \left[\bar{b} \right] Z(\bar{b})}{\left[\Phi(\bar{b}) - \Phi(\bar{a}) \right]} - \frac{\left[Z(\bar{b}) - Z(\bar{a}) \right]^2}{\left[\Phi(\bar{b}) - \Phi(\bar{a}) \right]^2} \right\} \sigma^2 = Var\{M\} \quad (10)$$

Example: A calibration certificate states that the mass of a stainless steel standard m_s of nominal value one kilogram is 1 000,000 325 g and that “the uncertainty of this value is 240 μg at the three standard deviation level ($k=3$)”. The standard uncertainty of the mass standard is then simply $u(m_s) = (240 \mu\text{g})/k = 80 \mu\text{g}$. This corresponds to a relative standard uncertainty $u(m_s)/m_s = 80 \times 10^{-9}$ the estimated variance is $u^2(m_s) = 6,4 \times 10^{-9} \text{ g}^2$.

So that $E\{M\} = 1\,000,000\,325 \text{ g}$, $\text{Var}\{M\} = 6,4 \times 10^{-9} \text{ g}^2$ and

$$a = E\{M\} - k\sqrt{\text{Var}\{M\}} = 1\,000,000\,085 \text{ g}$$

$$b = E\{M\} + k\sqrt{\text{Var}\{M\}} = 1\,000,000\,565 \text{ g}$$

$$\mu = E\{M\} = 1\,000,000\,325 \text{ g}$$

$$\sigma = \sqrt{\text{Var}\{M\}} \left[1 - \frac{2kZ(k)}{2\Phi(k)-1} \right]^{\frac{1}{2}} = 81,08832 \mu\text{g}$$

from (9) we have $f(m) = 0.01237 \frac{1}{\sqrt{2\pi}} e^{-\frac{(m-\mu)^2}{2\sigma^2}}$ with $a \leq m \leq b$

The maximum entropy is:

$$\text{MaxEnt} = - \int_a^b f(m) \ln(f(m)) dm = \ln(\sigma W \sqrt{2\pi}) + \frac{1}{2\sigma^2} \text{Var}\{M\}$$

$$\text{with } W = \Phi\left(\frac{b - E\{M\}}{\sigma}\right) - \Phi\left(\frac{a - E\{M\}}{\sigma}\right)$$

By considering the normal distribution in the interval $(-\infty, +\infty)$ we deduce $E\{M\} = \mu$, $\text{Var}\{M\} = \sigma^2$ and $\text{MaxEnt} = \ln(\sigma\sqrt{2\pi}e)$

2.2. Application of the MINCENT to a measurement model

For this application, we consider a general measurement process that treats multiple measurands and yields simultaneously multiple results or estimates.

We assume that all the relevant quantities involved in the process are interpreted or regarded as random variables.

In other terms all the sources of uncertainty are characterized by probability distribution functions, the form of which is assumed to either be known from measurements: the underlying probability distributions are sampled by means of repeated measurements and so they can be estimated through distribution-free influence, or unknown and so conjectured.

We classify all the involved quantities into two principal sets represented by row vectors:

1) output quantities \underline{Y} , in number of m

2) input quantities \underline{X} that comprise the rest of quantities, in number of n .

Let $(\underline{x}, \underline{y})$ the actual realizations of $(\underline{X}, \underline{Y})$ in a particular occasion, they represent a state of the measurement process in that occasion. The process has a set D of possible states $[(\underline{x}, \underline{y}) \in D]$ that identify the joint domain of the random variables $(\underline{X}, \underline{Y})$.

Further, often, it is possible in a measurement process to individuate mathematical and/or empirical models that link input and output quantities through functional relationships of type:

$$Y_i = g_i(X_1, \dots, X_n); i = 1, \dots, m \quad \text{with } m \ll n \quad (11)$$

In practical situations the transformation defined by (11) is differentiable and invertible.

The mutual behavior between the input quantities \underline{X} and \underline{Y} is statistically drawn by the joint probability density $f(\underline{x}, \underline{y})$ which can be written as:

$$f(\underline{x}, \underline{y}) = f(\underline{x})f(\underline{y} | \underline{X} = \underline{x}) \quad (12)$$

where $f(\underline{x})$ is the marginal joint density of the input quantities \underline{X} and $f(\underline{y} | \underline{X} = \underline{x})$, are the conditional joint densities of the output quantities \underline{Y} , given $\underline{X} = \underline{x}$.

In the Bayesian approach we have another important linkage between $f(\underline{x})$ and $f(\underline{x}, \underline{y})$, that is:

$$f(\underline{x} | \underline{Y} = \underline{y}) = c f(\underline{x}) f(\underline{y} | \underline{X} = \underline{x}) \quad (13)$$

where $f(\underline{x} | \underline{Y} = \underline{y})$ is the posterior joint density of the input quantities \underline{X} , given $\underline{Y} = \underline{y}$ and $f(\underline{x})$ is interpreted as the prior joint density which tells us what is known about \underline{X} without knowledge of \underline{Y} , while $f(\underline{y} | \underline{X} = \underline{x}) = \ell(\underline{x}, \underline{y})$, regarded as a function of \underline{x} for prefixed output values \underline{y} , represents the well-known likelihood; c is a "normalizing" constant necessary to ensure that the posterior joint density integrates, with respect to \underline{x} , to one.

We apply the Kullback's principle to the joint density function $f(\underline{x}, \underline{y})$ expressed by (12) and take into account the Bayesian relation (13) which discriminates between the prior densities of the input quantities \underline{X} and the conditional densities of the output quantities given the input ones.

To this end we introduce the joint cross entropy in the following manner:

$$S = \int_D f(\underline{x}, \underline{y}) \ln \frac{f(\underline{x}, \underline{y})}{f_0(\underline{x}, \underline{y})} \underline{dx} \underline{dy} = E \left\{ \ln \frac{f(\underline{X}, \underline{Y})}{f_0(\underline{X}, \underline{Y})} \right\} \quad (14)$$

where $\underline{dx} = dx_1 \cdots dx_n$, $\underline{dy} = dy_1 \cdots dy_m$ and $f_0(\underline{x}, \underline{y})$ is defined as an "invariant measure" function and it is a known "prior" knowledge on the measurement process.

Practical solution for MINCENT

By substituting eq. (12) into (14) and considering the domain D divisible into two sub-domains D_0 and D_1 corresponding respectively to the ranges of \underline{X} and of \underline{Y} we can write:

$$S = S_0 + E\{S_1(\underline{X})\} \quad (15)$$

where $S_0 = E \left\{ \ln \frac{f(\underline{X})}{f_0(\underline{X})} \right\} = \int_{D_0} f(\underline{x}) \ln \frac{f(\underline{x})}{f_0(\underline{x})} \underline{dx}$, having imposed the normalizing condition

$$\int_{D_1} f(\underline{y} | \underline{X} = \underline{x}) \underline{dy} = 1 \quad (16)$$

and

$$S_1(\underline{X}) = \int_{D_1} f(\underline{y} | \underline{X}) \ln \frac{f(\underline{y} | \underline{X})}{f_0(\underline{y} | \underline{X})} \underline{dy} \quad (17)$$

so that:

$$E\{S_1(\underline{X})\} = \int_{D_0} f(\underline{x}) \underline{dx} \int_{D_1} f(\underline{y} | \underline{X}) \ln \frac{f(\underline{y} | \underline{X})}{f_0(\underline{y} | \underline{X})} \underline{dy} \quad (18)$$

It is important to emphasize that, if $f_0(\underline{x})$ and $f_0(\underline{y} | \underline{x})$ are constants, S_0 and $S_1(\underline{X})$ coincide respectively with $-H_0$ and $-H_1(\underline{X})$, being H_0 the classical Jaynes' entropy referred to the density of the input quantities and $H_1(\underline{X})$ that one referred to the conditional density of the output quantities given the input quantities \underline{X} .

If $S_1(\underline{X}) = S_1$ would result independent of \underline{X} , by imposing the other normalizing condition:

$$\int_{D_0} f(\underline{x}) \underline{dx} = 1 \quad (19)$$

we should deduce the important additive property:

$$S = S_0 + S_1$$

Thus, in this case, the joint cross-entropy is equal to the cross-entropy referred to the input quantities plus the cross-entropy of conditional density of output quantities given the input ones.

If the constraints on the input quantities do not interfere with those ones of the output quantities in order to minimizing the general joint cross-entropy we can minimize before with respect to $f(\underline{y}|\underline{X} = \underline{x})$ and than to $f(\underline{x})$.

In a first step we minimize the expression $S_1(\underline{X})$, given by (17), subject either to the normalizing condition (16) and to other additional integral constraints of type:

$$E\{g_i(\underline{Y}, \underline{X})|\underline{X}\} = \int_{D_1} g_i(\underline{y}, \underline{X}) f(\underline{y}|\underline{X} = \underline{x}) d\underline{y} = g_i(\underline{X}) \quad i=1, \dots, t \quad (20)$$

where the functions $g_i(\underline{y}, \underline{X})$ and the values $g_i(\underline{X})$ are given, and in general they can depend on \underline{X} .

The minimum of cross-entropy $S_1(\underline{X})$ subject to constraints (16) and (20) is found through the well-known Lagrange variational method.

We must solve Euler's equation for the functional:

$$\int_{D_1} f(\underline{y}|\underline{X}) \left[\ln \frac{f(\underline{y}|\underline{X} = \underline{x})}{f_0(\underline{y}|\underline{X} = \underline{x})} - \lambda_0 - \sum_{i=1}^t \lambda_i g_i(\underline{y}, \underline{X}) \right] d\underline{y} \quad (21)$$

being λ_k ($k = 0, 1, \dots, t$) the well-known Lagrange's multipliers.

After simple manipulations we find the following solutions:

$$f(\underline{y}|\underline{X} = \underline{x}) = f_0(\underline{y}|\underline{X} = \underline{x}) e^{\lambda_0 - 1 + \sum_{i=1}^t \lambda_i g_i(\underline{y}, \underline{X})} \quad (22)$$

By substituting (22) into (17) we deduce the minimum of cross-entropy $S_1(\underline{X})$, that is:

$$S_{1\min}(\underline{X}) = \int_{D_1} f(\underline{y}|\underline{X} = \underline{x}) \left[\lambda_0 - 1 + \sum_{i=1}^t \lambda_i g_i(\underline{X}) \right] d\underline{y} \quad (23)$$

By recalling constraints (16) and (20) we deduce finally:

$$S_{1\min}(\underline{X}) = \lambda_0 - 1 + \sum_{i=1}^t \lambda_i g_i(\underline{X}) \quad (24)$$

Now, from (18), by taking into account (19), we obtain:

$$E\{S_{1\min}(\underline{X})\} = \lambda_0 - 1 + \sum_{i=1}^t \lambda_i E\{g_i(\underline{X})\} \quad (25)$$

Let us introduce (25) into (15) and we derive:

$$S = E \left\{ \ln \frac{f(\underline{X})}{f_0(\underline{X})} \right\} + \lambda_0 - 1 + \sum_{i=1}^t \lambda_i g_i(\underline{X}) \quad (26)$$

Now we consider, besides the normalization condition (19), additional constraints of type:

$$E\{w_i(X)\} = \int_{D_0} w_i(x) f(x) dx = w_i \quad i=1, \dots, v \quad (27)$$

where the functions $w_i(x)$ and the values w_i are known. The minimum of S with respect to $f(x)$ subject to constraints (19) and (27) is given, as previously, by introducing new Lagrange's multipliers λ_i ($i = 0, 1, \dots, v$).

We must solve a new Euler's equation for the following functional:

$$\int_{D_0} f(x) \left[\ln \frac{f(x)}{f_0(x)} + \lambda_0 - 1 + \sum_{i=1}^v \lambda_i g_i(x) - \lambda_0' - \sum_{i=1}^v \lambda_i' w_i(x) \right] dx \quad (28)$$

After simple manipulation, taking into account eq. (24) we obtain:

$$f(x) = f_0(x) e^{-S_{\text{imin}}(x) + \lambda_0' - 1 + \sum_{i=1}^v \lambda_i' w_i(x)} \quad (29)$$

Example: The repeated measurements

We repeat n times the measurements process on the same measurand according to the "repeatability conditions" established by the "Guide to the expression of uncertainty in measurement" (ISO 1993).

We may conveniently regard any set of measurement results $\underline{y} = (y_1, \dots, y_n)$ as the n -dimensional realization of an induced random vector $\underline{Y} = (Y_1, \dots, Y_n)$, which we can call, output vector.

Let us now assign the random variable X to the measurand and express its occasional realization as x , which we suppose constant during all the replications. We suppose that the conditional statistic parameters are given by:

$$\begin{cases} E\{Y_i|X=x\} = x \\ E\{(Y_i - x)^2|X=x\} = \sigma^2 \end{cases} \quad i=1, \dots, n \quad (30)$$

Let $f(x, \underline{y}) = f(x)f(\underline{y}|x)$ the unknown joint probability density of the measurand X and the random vector \underline{Y} , and $f_0(x, \underline{y}) = f_0(x)f_0(\underline{y}|x)$ the prior joint density, memorizing all the previous knowledge and experience.

Assuming the conditional independence of the output quantities Y_1, \dots, Y_n we can write:

$$\begin{cases} f(\underline{y}|x) = \prod_{i=1}^n f(y_i|x) \\ f_0(\underline{y}|x) = \prod_{i=1}^n f_0(y_i|x) \end{cases} \quad (31)$$

$f(y_i|x)$ and $f_0(y_i|x)$ being the common marginal conditional densities at the generic argument y_i , assuming $X=x$.

Taking into account eq. (31) we deduce the joint cross entropy in the form:

$$S(x) = n \int_{-\infty}^{+\infty} f(y|x) \ln \frac{f(y|x)}{f_0(y|x)} dy \quad (32)$$

Using Lagrange's multipliers method the MINCENT density is given by

$$f(y|x) = f_0(y|x) e^{-[\lambda_0 + \lambda_1 y + \lambda_2 (y-x)^2]} \quad (33)$$

where the Lagrange multipliers $\lambda_0, \lambda_1, \lambda_2$, are determined by using the given constraints. Assuming, for simplicity, a constant prior density $f_0(y|x) = k$, after simple manipulation, taking into account eq. (31) we obtain:

$$f(y|x) = \frac{1}{2\pi^{\frac{n}{2}} \sigma^n} e^{-\frac{\chi^2}{2}} \quad \text{where} \quad \chi^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (Y_i - x)^2 \quad (34)$$

Now we consider the Bayesian inference and introduce the posterior density of the measurand that is:

$$f_x(x|y_{\underline{0}}) = c f(x) f(y_{\underline{0}}|x) \quad (35)$$

where $y_{\underline{0}} = (y_0, \dots, y_{0_n})$ is the set of effective results at the particular occasion when $X=x$ and c is a convenient normalizing constant.

By introducing (34) into (35) and assuming $f(x)$ constant, we obtain:

$$f_x(x|y_{\underline{0}}) = c' e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_{0i} - x)^2} = c' e^{-\frac{1}{2\sigma^2} \left(nx^2 - 2 \sum_{i=1}^n y_{0i} x \right)} \quad (36)$$

where c' is the up-to-date normalizing constant.

The posterior expectation and variance of the measurand are given by:

$$E\{X|y_{\underline{0}}\} = c' \int_{-\infty}^{+\infty} x e^{-\frac{1}{2\sigma^2} \left(nx^2 - 2 \sum_{i=1}^n y_{0i} x \right)} dx = \frac{1}{n} \sum_{i=1}^n y_{0i} = \bar{y}_0, \quad Var\{X|y_{\underline{0}}\} = \frac{\sigma^2}{n} \quad (37)$$

References

1. BIPM, IEC, IFCC, ISO, IUPAC, and OIML, *Guide to the Expression of Uncertainty in Measurement*, 1995, ISBN 92-67-10188-9, Second Edition.
2. C.E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. J., vol.27, 1948.
3. J.N. Kapur, *Maximum-Entropy Models in Science and Engineering* John Wiley & Sons
4. M.G. Cox, M.P. Daiton and P.M. Harris, *Best Practice Guide No.6. Uncertainty and Statistical Modelling*. Technical report, National Physical Laboratory, Teddington, UK, 2001.

STOCHASTIC PROCESSES FOR MODELLING AND EVALUATING ATOMIC CLOCK BEHAVIOUR*

G. PANFILO

*Politecnico di Torino, Corso Duca degli Abruzzi 24, 10124 Torino, ITALY
E-mail: panfilo@tf.ien.it*

P. TAVELLA

*IEN "Galileo Ferraris", Strada delle Cacce 91, 10135 Torino, ITALY
E-mail: tavella@tf.ien.it*

C. ZUCCA

*Dip. di Matematica, Università di Torino, Via Carlo Alberto 10, 10123 Torino, ITALY
E-mail: zucca@dm.unito.it*

A model for atomic clock errors is given by the stochastic differential equation. The probability that the clock error does not exceed a limit of permissible error is also studied by means of the survival probability.

1. Introduction

The behavior of the atomic clocks is typically studied considering their phase and frequency error with respect to an ideal clock. This quantity is known to be well modelled by a stochastic process. We present a formal model of the clock phase error based on the theory of stochastic differential equations. In addition, we study the problem of the first passage time of the stochastic processes across two fixed constant boundaries. This topic is also known as survival probability between two absorbing barriers and it can be studied by means of the infinitesimal generator and by the use of simulation or numerical techniques. The two approaches are discussed and compared in order to show the importance of accurate numerical schemes to solve the equations.

We present the application of these results to the evaluation of the probability that the clock error exceeds the limit of permissible error at a certain time after synchronization. This aspect has interesting applications in satellite systems and also in the frame of the Mutual Recognition Arrangement where it may be important not only demonstrating that two standards are "equivalent", but also evaluating how long such equivalence can last. Results on experimental clock measures are presented.

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

2. Mathematical model for the atomic clock error

Typically the clock signal is affected by five random noises that are known in metrological literature as: White phase modulation, Flicker phase modulation, White frequency modulation (WFM), that produces a Brownian motion (BM, also called Wiener process) on phase, Flicker frequency modulation, Random walk frequency modulation (RWFM) which produces on integrated Brownian motion (IBM) on the phase. The presence of the different kinds of random noises and their entity is different from clock to clock. From experimental evidence it appears that on Cesium clock the predominant noises are the WFM and the RWFM that corresponds in mathematical language, to the Wiener process and to the Integrated Wiener process respectively on the phase error. In this paper we focus on the atomic clock model based on these two main noises. Denoting with $X_1(t)$ the time dependent evolution of the atomic clock phase error, it can be viewed as the solution of a dynamical system, written as a system of stochastic differential equations whose driving terms are two independent standard Wiener processes $\{W_1(t), t \geq 0\}$ and $\{W_2(t), t \geq 0\}$

$$\begin{cases} dX_1(t) = (X_2(t) + \mu_1)dt + \sigma_1 dW_1(t) \\ dX_2(t) = \mu_2 dt + \sigma_2 dW_2(t) \end{cases} \quad t \geq 0 \quad (1)$$

with initial conditions $X_1(0) = x$, $X_2(0) = y$, where μ_1 and μ_2 are constants that in the clock case represent what is generally referred to as the deterministic phenomena of the atomic clocks. In particular μ_1 is related to the constant initial frequency offset, while μ_2 represents what is generally indicated by frequency ageing or drift. The positive constants σ_1 and σ_2 represent the diffusion coefficients of the two noise components and give the intensity of each noise. The relationship between the diffusion coefficients and the more familiar Allan variance used in time metrology is known [1]. $W_1(t)$ is the Wiener noise acting on the phase X_1 driven by a white noise on the frequency. Since the Wiener process can be thought of as the integral of a white noise, the phase deviation X_1 is the integral of the frequency deviation affected by white noise. $W_2(t)$ represents the Wiener process on the frequency (the so called “Random Walk FM”) and gives an Integrated Wiener process on the phase.

Note that the initial conditions x and y should not be confused with the metrological notation where usually x represents the phase and y represents the frequency. Here we consider x and y as numbers but they can be considered as random variables (cf. Section 3). Note that X_1 represents the phase deviation, while the frequency deviation results to be \dot{X}_1 . The second component X_2 is

only a part of the clock frequency deviation, i.e. what is generally called the "random walk" component.

Since equation (1) is a bidimensional strictly linear stochastic differential equation, it is possible to obtain its solution in closed form [2,3]

$$\begin{cases} X_1(t) = x + (y + \mu_1)t + \mu_2 \frac{t^2}{2} + \sigma_1 W_1(t) + \sigma_2 \int_0^t W_2(s) ds \\ X_2(t) = y + \mu_2 t + \sigma_2 W_2(t) \end{cases} \quad (2)$$

This solution can be written in iterative form, that results very useful for further processing of the data and for simulations [3]. Let us consider now a fixed time interval $[0, T] \subset \mathbb{R}^+$ and a partition $0 \equiv t_0 < t_1 < \dots < t_N \equiv T$ equally spaced and let us denote the resulting discretization step with $\tau = t_{k+1} - t_k$, for each $k = 0, 1, \dots, N-1$. We can express the solution (2) at time t_{k+1} in terms of the position of the process at time t_k

$$\begin{cases} X_1(t_{k+1}) = X_1(t_k) + (X_2(t_k) + \mu_1) \tau + \mu_2 \frac{\tau^2}{2} + J_{k,1} \\ X_2(t_{k+1}) = X_2(t_k) + \mu_2 \tau + J_{k,2} \end{cases} \quad (3)$$

where $J_{k,1}$ and $J_{k,2}$ are the two components of the vector \mathbf{J}_k :

$$J_k = \begin{bmatrix} \sigma_1 (W_1(t_{k+1}) - W_1(t_k)) + \sigma_2 \int_{t_k}^{t_{k+1}} (W_2(s) - W_2(t_k)) ds \\ \sigma_2 (W_2(t_{k+1}) - W_2(t_k)) \end{bmatrix} \sim N \left(0, \begin{bmatrix} \sigma_1^2 \tau + \sigma_2^2 \frac{\tau^3}{3} & \sigma_2^2 \frac{\tau^2}{2} \\ \sigma_2^2 \frac{\tau^2}{2} & \sigma_2^2 \tau \end{bmatrix} \right).$$

The vector \mathbf{J}_k is the "innovation", i.e. the stochastic part that is added to build the process at the instant t_{k+1} . It depends only on the increments of the Wiener processes in the interval (t_k, t_{k+1}) . Such increment, $W(t_{k+1}) - W(t_k)$, is distributed as a Wiener process originated at the instant t_k on an interval of length τ , i.e. $N(0, \tau)$.

3. The survival probability and infinitesimal generator

Let $X(t)$ be a homogeneous stochastic Markov process starting in x_0 at time $t = 0$, $\{X(t), t \geq 0; X(0) = x_0\}$ and let $\tau_{(-n,m)}$ be the first passage time of $X(t)$ through two constant absorbing boundaries $-n$ and m

$$\tau_{(-n,m)} = \inf \{ t \geq 0 : X(t) \notin (-n, m); X(\theta) \in (-n, m) \forall \theta < t; X(0) = x_0 \}.$$

Let $f(t, x, B) = P(X_{t+s} \in B | X_s = x)$ be the transition probability density

function of X , i.e. the probability that the homogeneous Markov process that

starts in x at time s , reaches the set B at time $t+s$ and let $p(t,x)$ be the probability that the process starting in x at time $t = 0$ does not reach the boundaries in the time interval $(0,t)$

$$p(t,x) = P(\tau_{(-n,m)} \geq t | X_0 = x)$$

This function is called *survival probability*. The application of the survival probability is very important in our applications because we can interpret it as the probability that the clock error has not exceeded the limit of permissible errors at a certain time after synchronization.

The infinitesimal generator is defined [4,5] as the operator A s.t.

$$Ag(x) = \lim_{t \rightarrow 0^+} \frac{1}{t} [T_t g(x) - g(x)]$$

where g is a bounded function and T_t is an operator defined as

$$T_t g(x) = \int g(y) f(t,x,dy) = E^x[g(X_t)]$$

$Ag(x)$ can be interpreted as the mean infinitesimal rate of change of $g(X_s)$ in case $X_s = x$. It is possible to prove [4] that there is a strong link between the stochastic differential equation and the partial differential equations via the infinitesimal generator. Indeed, if we consider the m -dimensional general stochastic differential equation:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

where $\mu: \mathfrak{R}^m \rightarrow \mathfrak{R}^m$, $\sigma: \mathfrak{R}^m \rightarrow \mathfrak{R}^{m \times m}$ are measurable functions, $x \in \mathfrak{R}^m$, the following expression provides the infinitesimal generator:

$$L_t = \frac{1}{2} \sum_{i,j=1}^m (\sigma(x)\sigma^T(x))_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m \mu_i(x) \frac{\partial}{\partial x_i} \quad (4)$$

where σ^T is the transposed matrix.

The infinitesimal generator is very important in the study of the stochastic processes since it allows to determine stochastic quantities solving partial differential equations.

For example, the infinitesimal generator may be used to study the survival probability by means of the following partial differential equation:

$$\begin{cases} L_t p(t,x) = \frac{\partial p(t,x)}{\partial t} & \text{on } [0,T] \times D \\ p(0,x) = 1_D \\ p(t,x) = 0 & \text{on } [0,T] \times \partial D \end{cases} \quad (5)$$

where the function p , solution of (5), is the survival probability defined on the time domain $[0,T]$ and on the space domain D .

The same partial differential equation applied to the transition probability density function leads to the Kolmogorov backward equation.

3.1. Examples

In this section we study to the clock model introduced in Section 2 using the PDE (5). In particular, if we consider the bidimensional model (1) with $\mu_1 = 0$, $\mu_2 = 0$ without loosing generality, we get that the infinitesimal generator related to this stochastic process is:

$$L_t = y \frac{\partial}{\partial x} + \frac{1}{2} \sigma_2^2 \frac{\partial^2}{\partial y^2} + \frac{1}{2} \sigma_1^2 \frac{\partial^2}{\partial x^2}$$

obtained by (4) using:

$$\mu = \begin{pmatrix} y \\ 0 \end{pmatrix} \sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \sigma \sigma^T = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}.$$

As a first approach, we focus our attention on the separate studies of the simple BM or of the IBM. The BM can be defined via the following one-dimensional stochastic differential equation:

$$dX_t = \mu dt + \sigma dW_t, \quad (6)$$

where now $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ are constants. The infinitesimal generator of (6) results to be:

$$L_t = \mu \frac{\partial}{\partial x} + \sigma^2 \frac{1}{2} \frac{\partial^2}{\partial x^2}. \quad (7)$$

Note that this is the generator of the classical heat equation. In this case it is possible to derive the closed form solution of the PDE (5) i.e. the survival probability of the BM that is [6]:

$$p(t, x) = \frac{1}{\sigma \sqrt{2\pi t}} \sum_{k=-\infty}^{\infty} \left[\exp \left\{ \frac{\mu x'_k}{\sigma^2} - \frac{(x - x'_k - \mu t)}{2\sigma^2 t} \right\} - \exp \left\{ \frac{\mu x''_k}{\sigma^2} - \frac{(x - x''_k - \mu t)}{2\sigma^2 t} \right\} \right]$$

where $x'_k = 2k(m+n)$, $x''_k = 2m - x'_k$ with $k = 0, \pm 1, \pm 2, \dots$

On the other hand, the IBM is defined by the SDE (1) considering $\sigma_1 = 0$, $\mu_1 = 0$. Without loosing generality, as a first study case, we set $\mu_2 = 0$ and the infinitesimal generator results to be:

$$L_t = y \frac{\partial}{\partial x} + \frac{1}{2} \sigma_2^2 \frac{\partial^2}{\partial y^2}. \quad (7)$$

For the IBM the only known analytical result concerns the mean value of the first passage time across two fixed constant boundaries [7,8]. The analytical expression of the survival probability is not known, then we need numerical methods or simulations.

In the following we present two numerical methods that allows to obtain an estimation of the survival probability [9,10].

4. Finite difference and Monte Carlo methods

One numerical method, that makes use of the SDE is a Monte Carlo method. It consists in the simulation of a given number of realizations of the considered stochastic process and in the identification of the first passage time $\tau_{(-n,m)}$ of each simulated trajectory across the two constant boundaries $-n$ and m . From the statistics of $\tau_{(-n,m)}$ it is possible to estimate several quantities such as the survival probability.

The second numerical method makes use of the PDE (5) with the aid of the finite difference method. This method is based on the discretization of the domain and on the substitution of the continuous derivatives with discrete approximations

$$\begin{aligned}\frac{\partial p}{\partial x} &= \frac{p_{i+1,j} - p_{i,j}}{h} + o(h) \\ \frac{\partial^2 p}{\partial x^2} &= \frac{p_{i+1,j} - 2p_{i,j} + p_{i-1,j}}{h^2} + o(h^2).\end{aligned}$$

Hence, in this case the survival probability is estimated on a discrete grid.

4.1. Application to the Brownian motion

In this subsection we consider the application of the two numerical methods to the simple case of a BM (6) with $\mu = 0$.

In order to apply the Monte Carlo method, we simulate $N = 10^5$ trajectories, using the corresponding iterative form, obtained as solution of (6) [2],

$$X(t_{k+1}) = X(t_k) + \sigma (W(t_{k+1}) - W(t_k))$$

with discretization step $\tau = 0.01$ and diffusion parameter $\sigma = 1$, and we detect the first passage time of each trajectory across two fixed constant boundaries $-n$ and m with $m = n = 1$.

On the other side, the PDE (5) with the BM generator (7) is written as

$$\begin{cases} \frac{1}{2} \frac{\partial^2 p(t,x)}{\partial x^2} \sigma^2 = \frac{\partial p(t,x)}{\partial t} & t \in [0, T] \\ p(t, -n) = 0 \\ p(t, m) = 0 \\ p(0, x) = 1_D \end{cases} \quad x \in D = [-n, m]$$

and the application of the finite difference method leads to the following discretization scheme in terms of the survival probability $p(t, x)$

$$\begin{cases} \frac{1}{2} \frac{p_{i,j-1} - 2p_{i,j} + p_{i,j+1}}{h_x^2} \sigma^2 = \frac{p_{i,j} - p_{i-1,j}}{h_t} \\ p(t,1) = 0 \\ p(t,-1) = 0 \\ p(0,x) = 1_D. \end{cases}$$

This scheme is implemented with discretization steps $h_x = 0.1$, $h_t = 0.01$.

In figure 1 the survival probabilities obtained with the analytical solution, the Monte Carlo method and the finite differences method are compared but, since the three estimated curves result superimposed, the analytical and numerical solutions perfectly agree.

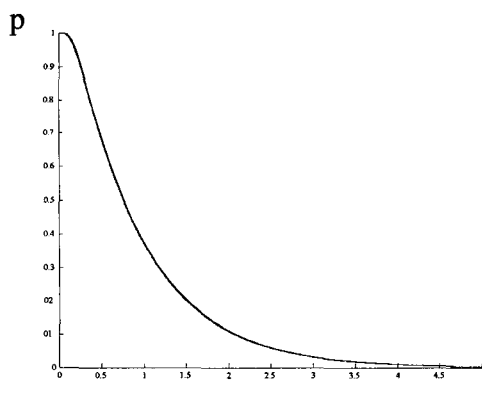


Figure 1. Survival probabilities obtained with the analytical solution , the Monte Carlo method and the finite differences method in case of BM (all lines are superimposed).

4.2. Application to the integrated Brownian motion

In this section we focus our attention on the IBM case in the case $\mu_2 = 0$. In order to apply the Monte Carlo method, we simulate $N = 10^5$ trajectories, using the corresponding iterative form obtained from (3) with $\mu_1 = \sigma_1 = 0$

$$\begin{cases} X_1(t_{k+1}) = X_1(t_k) + X_2(t_k)\tau + \sigma_2 \int_{t_k}^{t_{k+1}} W(s)ds \\ X_2(t_{k+1}) = X_2(t_k) + \sigma_2 (W(t_{k+1}) - W(t_k)) \end{cases}$$

with discretization step $\tau = 0.01$ and diffusion parameter $\sigma_2 = 1$, and we detect the first passage time of each trajectory across two fixed constant boundaries $-n$ and m with $m = n = 1$.

The survival probability p for the IBM is defined by the PDE (5) with IBM generator (8)

$$\begin{cases} \frac{\partial p}{\partial t} = y \frac{\partial p}{\partial x} + \sigma^2 \frac{1}{2} \frac{\partial^2 p}{\partial y^2} & t \in [0, T] \\ p(t, -n, y) = 0 & y > 0 \\ p(t, m, y) = 0 & y < 0 \\ p(t, x, -\infty) = 0 \\ p(t, x, \infty) = 0 \\ p(0, x, y) = 1_D. \end{cases} \quad (x, y) \in D = \{[-n, m] \times \mathbb{R}\}$$

Here we notice that the unknown $p(t, x, y)$ depends on the time t , the initial values x and y . The application of the implicit difference method leads to the following discretization scheme according with the method presented in [7]

$$\begin{cases} \frac{1}{2} \sigma^2 \frac{p_{i,j-1,k+1} - 2p_{i,j,k+1} + p_{i,j+1,k+1}}{h_y^2} + y_j \frac{p_{i+1,j,k+1} - p_{i,j,k+1}}{h_x} = \frac{p_{i,j,k+1} - p_{i,j,k}}{h_t} \\ \quad i = 0, \dots, n_x \quad j = 1, \dots, n_y \quad k = 0, \dots, n_t - 1 \\ \frac{1}{2} \sigma^2 \frac{p_{i,j-1,k+1} - 2p_{i,j,k+1} + p_{i,j+1,k+1}}{h_y^2} = \frac{p_{i,j,k+1} - p_{i,j,k}}{h_t} \\ \quad i = 1, \dots, n_x \quad j = 0 \quad k = 0, \dots, n_t - 1 \\ \frac{1}{2} \sigma^2 \frac{p_{i+1,j-1,k+1} - 2p_{i+1,j,k+1} + p_{i+1,j+1,k+1}}{h_y^2} + y_j \frac{p_{i+1,j,k+1} - p_{i,j,k+1}}{h_x} = \frac{p_{i+1,j,k+1} - p_{i+1,j,k}}{h_t} \\ \quad i = 0, \dots, n_x \quad j = -n_y, \dots, -1 \quad k = 0, \dots, n_t - 1 \end{cases}.$$

We then get a linear system of $2(n_x + 1)n_y n_t + n_x n_t$ variables. The matrix involved is sparse and, even if it is not symmetric, it is a band matrix.

In figure 2 the survival probabilities obtained via the Monte Carlo method and the finite differences method implemented with discretization steps $h_x = 0.04$, $h_y = 0.5$ and $h_t = 0.05$ or $h_t = 0.01$ are compared.

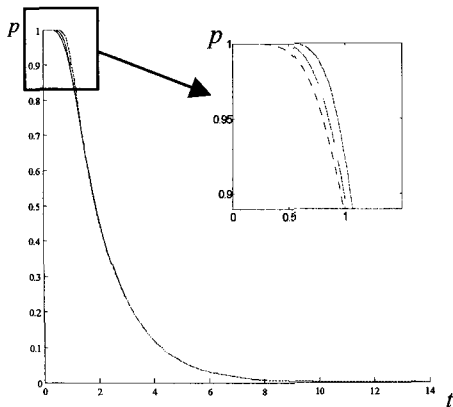


Figure 2. Survival probabilities obtained via the Monte Carlo method (solid line) and the finite differences method implemented with discretization steps $h_x = 0.04$, $h_y = 0.5$ and $h_t = 0.05$ (dotted line) or $h_t = 0.01$ (dashed line) in case of IBM.

We can notice that the three curves agree to a large extent but some difficulties arise in managing very small discretization steps since it leads to the handling of very large matrices.

5. Applications to the atomic clock error

The survival probability is an interesting quantity from the applicative point of view since it tell us how long we can wait, before re-synchronizing the clock, to be sure, at a certain percentage level, that the error has not crossed two fixed constant boundaries. In fact if we evaluate the abscissa t corresponding to a survival probability, for example $prob = 0.95$, we can estimate the time t that we should wait before the process crosses one of the boundaries with probability 0.95.

In Table 1 we report the time t (in days) that can elapse before the clock error exceeds the boundaries for different percentage levels and different values of the boundaries $m = n$. These results were obtained for a BM with drift $\mu = 0$ and variance coefficient $\sigma = 3.05$, i.e. the noise level of a typical cesium clock.

Table 1. Values of survival time t (in days) corresponding to different percentage levels of the survival probability and to different choices of the boundaries $m = n$ (in nanoseconds) (BM case with $\mu = 0$ and $\sigma = 3.05$)

$m \backslash prob$	90%	95%	99%
---------------------	-----	-----	-----

10	2.8	2	1.3
30	25.5	19	12.1
50	69.6	53.2	34.1
100	279	213	136
300	2500	1920	1220
500	6990	5340	3410

For example, fixing the maximum error allowed at ± 10 ns, from Table 1 we can deduce that we can wait two days being sure at 95% that the error has not yet exceed the permissible error.

Table 2 concerns a similar analysis of an IBM with drift $\mu_2 = 0$ and variance coefficient $\sigma_2 = 1$. Also in this case we choose the parameters of a typical Cesium clock. Fixing the maximum error allowed at ± 10 ns, from Table 2 we can deduce that we can wait four days being sure at 95% that the error has not exceed the permissible error.

Table 2. Values of time t (in days) corresponding to different percentage levels of the survival probability (IBM case with $\mu_2 = 0$ and $\sigma_2 = 1$)

$m \backslash prob$	90%	95%	99%
10	4.5	4	3.33

6. Conclusion

The paper describes the model of the atomic clock error using the stochastic differential equation. In addition, the related properties such as the Partial Differential Equation describing the survival probability have been studied because they can help in understanding and predicting clock error in many different metrological applications. In fact, the study of the survival probability indicates how often a clock has to be re-synchronized not to exceed a limit of permissible error. The accurate study of these properties are of great interest also for the applications to satellite systems and to the Mutual Recognition Arrangement.

References

1. J. W. Chaffee, *IEEE Transactions on Ultrasonic, Ferroelectrics, and Frequency Control* **34**, 6 (1987).

2. P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equation*. Berlino Heidelberg, Springer-Verlag (1992).
3. L. Galleani, L. Sacerdote, P. Tavella and C. Zucca, *Metrologia* **40**, 257 (2003).
4. P. Baldi, *Equazioni differenziali stocastiche applicazioni*. U.M.I., Bologna, Pitagora Editrice (2000).
5. S. Karlin and H. M. Taylor, *A second course in stochastic processes*. New York, Academic Press (1981).
6. D. R. Cox and H. D. Miller, *The theory of stochastic processes*. London, Chapman and Hall (1972).
7. J. N. Franklin and E. R. Rodemich, *S.I.A.M. J. Numer. Anal.* **5**, 680 (1968).
8. J. Masoliver and J. M. Porrà, *Phys. Rev. Lett.* **75**, 189 (1995).
9. W. H. Press et al., *Numerical Recipes in Fortran*. 2 ed. Camb. Univ. Press (1995).
10. G. Monegato, *Calcolo numerico*. Torino, Levrotto & Bella (1989) (In Italian).

COMPOUND-MODELLING OF METROLOGICAL DATA SERIES*

F. PAVESE

CNR, Istituto di Metrologia “G.Colonnetti”, strada delle Cacce 73, 10135Torino,
Italy. E-mail: f.pavese@imgc.cnr.it

In metrology the measurements on the same standard are usually repeated several times in each Laboratory. Repeating measurements is also the idea behind making inter-comparisons in different Laboratories of standards of a physical or chemical quantity. Often, whether these results can, or not, be considered as repeated measurements is not obvious, so that the statistical treatment of the data as they are repeated data can bring to misleading results. The paper reviews the use of two classes of methods keeping track of the fact that the data are collected in series: a) those considering a class of regression models able to accommodate both the commonality of all series and the specificity of each series; b) those using the mixture probability model for describing the pooled statistical distribution when the data series are provided in a form representing their statistical variability. Some problems related to the uncertainty estimate of the latter are also introduced.

1. Introduction

An essential method in metrology consists in repeating in each Laboratory several times the measurements on the same standard and this basic method also applies to inter-comparisons in different Laboratories of standards of a physical or chemical quantity.

In both cases, a basic question in order to decide the statistical treatment of the overall data is whether they can, or not, be considered as repeated measurements. Often the answer is not obvious, and consequently the statistical treatment of the data as a compound—data fusion—based only on the *a priori* assumption that they are repeated data, can bring to misleading results.

Particularly important in metrology are the mathematical and statistical methods that are able to take into account the fact that the data are collected in series. Using models that are able to take into account both the specificity and the commonality of the data series can better perform the operation of collation, often also called “data fusion”. Then, statistical tests can be performed *a posteriori* on rational bases on the estimated quantities to verify the repeated-measurement hypothesis. The paper discusses the use of compound modelling:

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

- a) when the data depends on one independent variable in a range. They need the fitting of a regression model function, which needs to accommodate both the commonality of all series of data and the specificity of each series;
- b) when each series of data are not provided but a “summary” form is. The summary represents the data statistical variability, e.g., when the probability density function (pdf) is provided as the overall information concerning each series of data. In these cases, the use of a mixture probability model for the pooled statistical distribution does not require the repeated-measurement hypothesis.

These methods are summarised in the paper and can be applied either to data series taken in a single Laboratory or to inter-comparison data.

2. Data in Series

Several measurements are taken of the same standard, yielding to data:

$$x_1^{(1)}, \dots, x_{N_1}^{(1)} \quad (1)$$

where each $x_n^{(1)}$ is affected by a random error $\varepsilon_n^{(1)}$. Several series of these data exist (e.g., measurements taken at different times, in different Labs, in different conditions, ...):

$$\begin{array}{c} x_1^{(m)}, \dots, x_{N_m}^{(m)} \\ \dots \\ x_1^{l(M)}, \dots, x_{N_M}^{(M)} \end{array} \quad (2)$$

The metrologist has to devise a summary of the information conveyed in the $\sum_{m=1}^M N_m$ data (2). The procedure bringing to the choice of the summary comprises, at least:

- a) the identification of the measurand(s) according to the purpose of the measurements;
- b) the choice of a statistical treatment well suited to obtain the expectation value of the measurand(s);
- c) the estimate of the uncertainty associated to the information represented by the overall data (2).

The importance—and sometimes the difficulty—of step a) is sometimes underestimated. The identification can be difficult or even controversial, arising from the purpose of the measurements and/or of the data analysis. In all instances, it

determines the type of mathematical modelling and of statistical treatment that is more appropriate for the subsequent steps.

Let us use for each line of (2) the following

DEFINITION: The series is a set of repeated measurements (usually homoscedastic or with $\text{var}(\varepsilon_i) = \text{const}$).

2.1. Types of standards

Different types of standards require different answers to the question a) about the meaning of the series of data. The meaning is different for two general classes defined in [1,2].

a) Set of artefact standards

DEFINITION: An artefact standard is called a measurement device, a “natural value” of which *does not* exist.

Examples are the mass or the length of a piece of metal. Each artefact carries its own value of the physical or chemical quantity, which can be estimated only by calculation and/or through comparison with others. From a statistical point of view, having a set of M artefact standards, each m -th artefact—one line in (2)—is a *different measurand* and each measurement series $x^{(m)}$ pertains to a different population, i.e. to a different random variable $X^{(m)}$. The measurement series can be obtained either in one Laboratory (M local standards) or in a set of M Laboratories (inter-comparison of M standards of different Laboratories).

b) Single standard

A definition of a single standard applies to several cases:

- i) each *single standard* in a Laboratory, where M series of measurements are performed on it, typically at subsequent times (an artefact also pertains to this category as far as it is considered individually and M *infra*-Laboratory evaluations are concerned);
- ii) *single travelling standard* used in an inter-comparison of M Laboratories;
- iii) set of M realisations of the same state of a physical quantity, a *natural value* of which exists, representing the *single standard* (concerning either *infra*- or *inter*-Laboratory measurements).

In all these cases, from a metrological and statistical viewpoint, there is only one measurand and all data $x^{(m)}$ pertain to the *same population*, i.e., to a single random variable, Q , while still the data is subdivided in series.

2.2. One random variable versus several

There is a basic difference in the statistical reasoning that must be applied to the data series, depending on whether they deal with one or with several (two or more) random variables.

a) *Several variables* (as in class a) standards), one for each m -th data series.

The synthetic information concerning the set of variables $X^{(m)}$ involves a summary statistic, typically a mean value, E .

The new random variable E is at a different (upper) level with respect to the input variables $X^{(m)}$, i.e., a hierarchical reasoning is introduced [6]. The uncertainty evaluation involves products of the probability distribution functions $\mathcal{F}^{(m)}$ associated to the $X^{(m)}$ —convolution in the linear case. In fact, there is uncertainty propagation in the sense of GUM [3].

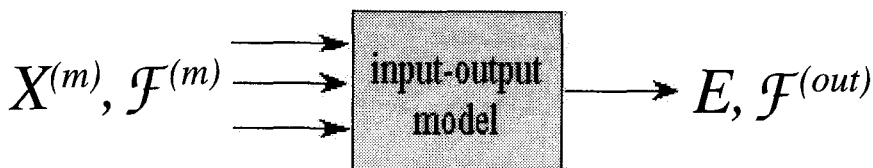


Figure 1. Typical GUM model.

b) *One variable* (as in class b) standards), M data series.

The operation of compounding the $x_n^{(m)}$ data is called data pooling, also referred to as “data fusion” (of the series) [12]. Since the variable Q is common to all data, no input-output model (Fig. 1) of the measurements exists and is needed. A single probability distribution function \mathcal{F} is associated to Q , as the union (summation) of the distribution functions $\mathcal{F}^{(m)}$ of each series of pooled data.

2.3. Several data series for a single random variable

The opposite attitude of overlooking the occurrence of a single standard and, thus, of always considering each m -th series as pertaining to a separate random variable, is to assume all data series as pooled into a single series of $\sum_{m=1}^M n_m$ data, overlooking—and losing—the information related to the possible series individuality.

On the contrary, series individuality is information *additional* to the one provided by their commonality, which should be analysed in order to allow the metrologist to reply to important questions (see Section 2.1.b):

Case i) (one standard in a Laboratory)— Is the standard stable in time? Is it reproducible? Can the data of subsequent series taken in the same Laboratories be considered as repeated measurements?

Case ii) (single travelling standard of an inter-comparison)— Is the standard stable in time? Is it reproducible? Can the data of all the Laboratories participating to the inter-comparison be considered as repeated measurements? (similar to *Case a)*

Case iii) (physical state realisation, either in one or in several Laboratories)— Are the realisations of the standard equally well implementing the physical state? Can all realisations (data series) be considered as repeated measurements?

2.4. Compound modelling techniques

The models keeping explicit both the individuality of each series and the commonality of all data will be referred to in the following as “compound models” and two techniques using them are briefly discussed in the following:

- the input-output model: $y_i^{(m)} = f(x_i^{(m)})$ for data series;
- the model of the statistical distribution function for single-standard problems.

A more comprehensive discussion can be found in [2, 4–6]. The basic advantage of their use is that, after having determined the parameters of the compound model, statistical tests can then be applied to the series-specific parameters to check if they are significant or not. If not, all the data can be considered as repeated measurements: this decision is now based on a statistical analysis of each series and not an *a priori* unverified assumption. If all or some of the series are significantly specific, the metrologist has a sound basis for further inference and data analysis.

3. Compound-modelling of Data Series in Regression

In regression of data collected in series, the so-called “Least Squares Method with Fixed Effect” (LSMFE) uses a model written, in general, as:

$$y_i^{(m)} = \text{Common}(x_i) + \text{Specific}(x_i^{(m)}) \quad (3)$$

with $m = 1, \dots, M$ and $i = 1, \dots, n_m$. For example, with a polynomial model for both parts, and of degree one for the *Specific* one, the regression model is:

$$g(x) = \sum_{u=0}^U a_u x^u + \sum_{m=1}^M \delta_m (t_m + b_m x) \quad (4)$$

where $\delta_m = 1$ if x is in the m -th series, otherwise is zero and the a_u , the t_m and the b_m are regression parameters to be determined. The $(U+2M+1)$ regression parameters are independent only if all t_m sum to zero (or other fixed prescribed constant). If the r -th series is taken as the "reference", the r index is skipped in the second summation and only $(U+2M-1)$ parameters are to be computed. A statistical significance test can then be applied to the t_m and b_m .

In figure 2 the simplest possible case of series differing only by a translation is reported, less visually evident when Type A standard uncertainties are large, as in figure 3. More applications can be found in [4, 5].

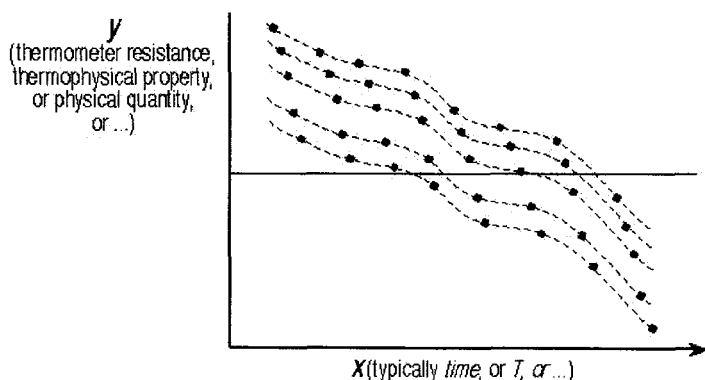


Figure 2. Series differing by a translation: *Specific = const* (simplest example).

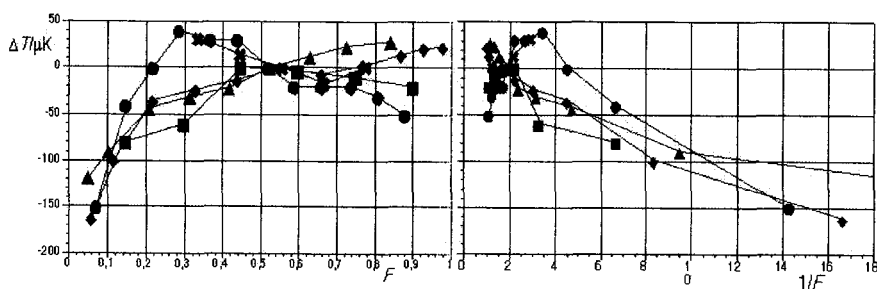


Figure 3. Example of series differing by a translation with predominant Type A standard uncertainties.

4. Compound Probability-Distribution Models for Single-Standard Data Series

Let us now consider the problem of describing data variability in an inter-comparison (IC) of standards. The n -th participant's data series is of the form (1) and (2) represents the overall input data to the IC, one series for each participant (local data). However, in the case of an IC, the local data are normally summarised to provide three items, representing the overall statistical properties of the local standard:

$$\{(Y_n \pm U_n), \mathcal{F}_n\}, \quad n = 1, \dots, N \quad (5)$$

where Y_n is the value assigned to the local standard and, in the probabilistic sense, is the location parameter of the local population and U_n is the expanded uncertainty associated to Y_n generally set at 95% confidence level. \mathcal{F}_n is the statistical distribution associated to the local standard to identify the population: at least the general class of models assumed for it should be provided [3].

Consider, however, an IC of standards based on different realizations of the same physical state (*class b*), *case iii*) in Section 2.1, e.g., temperature, pressure. For a full treatment of this case see [6]; for a comparison with the case of artefacts see [7]. Data are regarded as sampled from a single stochastic variable, Q . Therefore, for the sample $(x_{1n}, \dots, x_{jn}, \dots, x_{mn})$ of every n -th participant

$$x_{jn} \approx Q \quad \forall n \quad (6)$$

Since all data series $(x_{1n}, \dots, x_{jn}, \dots, x_{mn})$ pertain to the same population, they should be considered, at least *a priori*, as homogeneous and, consequently, it is appropriate to pool the local samples into a single sample (of total size $\sum_{m=1}^M n_m$)

and to estimate an expected value of the resulting super-population. Summary operations in each Laboratory, i.e., the definition of a new stochastic variable Y_n combining the local samples to estimate the local summary parameters and its U_n should be avoided since they would improperly introduce a hierarchy.

Q is distributed according to a distribution function \mathcal{F} , while each local probability distribution \mathcal{F}_n of the n -th population is sufficient to characterize the local stochastic variability of the n -th population. $\mathcal{A}(\mathcal{A})$ can be defined as the mixture of density functions, or mixture model of the super-population [6]:

$$f(\mathbf{x}, \mathcal{A}) = \sum_{n=1}^N \pi_n f_n(\mathbf{x}; \mathcal{A}_n), \quad \sum_{n=1}^N \pi_n = 1, \quad n = 1, \dots, N \quad (7)$$

each $f(x; A)$ is the density function of \mathcal{F}_n ; $\pi_n > 0$, where π_n are the proportions. The distribution function \mathcal{F} is the compound distribution of the N components \mathcal{F}_n .

The description of the overall data variability with a single pdf f does not require further assumptions, except those embedded in the identification of the local density functions f_n . Thus, (7) directly represents the overall variability of the physical state realisation, as resulting from the IC. An example of a bimodal distribution is shown in Fig. 4 [8]. In this case the metrologist has guidance for further inference about the data analysis [12]. In less evident cases, he can apply statistical tests to decide if the distribution is significantly different from a uni-modal one or from a given class of distributions, e.g., Normal, \mathcal{N} .

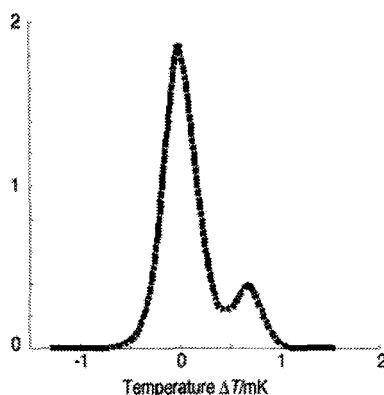


Figure 4. Mixture density distribution for a 1984 inter-comparison (argon).

In any instance, from the mixture distribution function, the IC summaries can be obtained in a straightforward way. The value called Reference Value (RV) in the Key Comparisons (KC) of the MRA [9] is the distribution expected value, the numerical value r :

$$r = E(\mathcal{F}) \quad (8)$$

The computation of (8) can be performed analytically or numerically.

The value of the Degree of Equivalence (DoE) of the KC, $a_{n,r}$, for the n -th participant is:

$$a_{n,r} = E(\mathcal{F}_n) - r \quad (9)$$

4.1. Evaluation of a KC uncertainty

The discussion of the problem of evaluating the uncertainty of a KC has so far been essentially associated to the discussion of the Reference Value (RV),

considered as a necessary step to evaluate the uncertainty of the DoE required by the MRA. This might not always be appropriate. Here the possible problems are summarised, which deserve a much more comprehensive analysis.

4.2.1 The RV as a deterministic parameter (DRV)

The RV is not necessarily a random variable. A Deterministic Reference Value (DRV) can be defined, which is a numerical value, r , with no uncertainty associated to it. Irrespective to the method used to obtain it, the DRV is a purely stipulated value, as opposed to a statistically-generated one. Its purpose is limited to the computation of the DoE, where it determines the deviation of the measurand value of every participant from the RV.

This kind of RV definition is not in contrast with the MRA requirements and can have the advantage to avoid the problems of a critical choice of the summary statistics, when its meaning or use is not essential to the use of the inter-comparison results. In addition, it avoids the problems arising from the correlation of the RV with the participants' values, as no uncertainty is associated to the DRV. Only the Degree of Equivalence (DoE) has an associated uncertainty, as required by MRA [9]. A DRV has been preferred for temperature key comparisons CCT K2 and CCT K4 [10,11].

In the case a DRV is used, the estimate of the uncertainty for the DoE of the n -th participant is simply:

$$u_{n,r} = u_n = u(x_n) \quad (10)$$

This would also be the uncertainty to associate to the DoE in (9).

4.2.2 The RV with an associated uncertainty

However, most KCs *do* associate an uncertainty to the RV. When several local random variables are defined (class a) standards), the RV is a random variable, whose uncertainty is estimated according the usual statistical methods.

When all the local data series pertain to the same Q (class b) standards) there are at least two possible ways to estimate the uncertainty:

- from the statistical variability of Q , as represented by \mathcal{F} . Usually, the second moment of the distribution is taken as the measure of u_r .
- from the set of the local uncertainties $\{u_n\}$ provided by the participants.

The first way to define the uncertainty certainly provides the usual description of the to-date variability in the current knowledge of the physical state (e.g., temperature value at the thermodynamic state called triple point), if the KC

exercise is comprehensive, e.g., comprises most worldwide-available realizations of the physical state. However, it is also evident that, by accepting all results as valid and with no screening of the participants' quality —prevented by the MRA,— the second moment of \mathcal{F} can be dominated by the uncertainty of the participants of lesser quality. This situation becomes extreme for an inter-comparison including only one very accurate and one very inaccurate Laboratory.

Therefore, the second way to define the KC uncertainty might be more suitable to picture the actual average quality of the participants, especially in top-accuracy metrology, where the u_n are dominated by Type-B uncertainties and consequently u_n represent more the overall quality of each Laboratory than the variability of its x_r .

The set $\{u_n\}$ of local uncertainties needs be summarised. In the case of a KC with many participants, the median, acting like a kind of “majority rule” of the average participant measurement capability, could be a suitable choice:

$$u_r = \text{median}(u_n), \quad n = 1 \dots N, \quad \mathcal{F}(u_r) = \mathcal{F}(\text{median}) \quad (11)$$

On the other hand, this would not be the case for an inter-comparison with only two participants. Alternatively, the metrologists could resort to some kind of consensus value for the KC uncertainty, as opposed to one purely based on sound statistical reasoning.

4.2.3 No RV defined

It is a peculiar case allowed by the MRA in special cases. Obviously, only the DoE between each pair of Laboratories can be defined and its uncertainty evaluation —and the KC uncertainty evaluation— yet requires a different kind of reasoning. However, does the MRA really allow for a non-uniform way to evaluate the uncertainty of different KC's —hence of the international equivalence— only because of a different choice of the RV?

5. Conclusions

The paper introduces the common problem in measurement science arising when the treatment of data series needs to take into account both their common and their series-specific characteristics. These problems find a different solution in methods suitable for different types of data series.

Two types of metrological data series have been discussed, for which in recent years mathematical and statistical tools have been adapted to the

metrological field. The first is when the data series need to be fitted to a single regression model in a range of values of the independent variable (e.g., data from different sources or from an unstable standard): the LSMFE solves the problem in a suitable way. The second is when each data series is summarized according to its statistical variability and all data series pertain to the same population, though arising from different sources: the use of a single (mixture) distribution function to describe the overall variability of the pooled data solves the problem and provides a straightforward definition of the required statistical parameters. In the latter case, the peculiar problems of defining an uncertainty for the overall data is introduced and will deserve further studies.

References

1. P.Ciarlini, F.Pavese, Some basic questions in the treatment of intercomparison data, (St.Petersburg Workshop, June 2002) *Izmeritel'naya Tekhnika*, 2003, N.4 (April), 68–71.
2. P.Ciarlini, M.G.Cox, F.Pavese, G.Regoliosi, The use of a mixture of probability distributions in temperature inter-laboratory comparisons, *Metrologia*, in press.
3. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML, *Guide to the Expression of Uncertainty in Measurement*, International Standards Institution, Second Edition, 1995.
4. F.Pavese, P.Ciarlini, Accurate modelling of translational bias and its application to the reduction of thermodynamic data series, *Metrologia*, 1990, 27, 145–152.
5. D.Ichim, P.Ciarlini, F.Pavese, A. Premoli, M.L. Rastello, Comparison of LS techniques for the linear approximation of data affected by heteroschedastic errors in both variables with uncertainty estimation, *This Book*.
6. P.Ciarlini, F.Pavese, G.Regoliosi, A bootstrap algorithm for mixture models and interval data in inter-comparisons. (Huddersfield Conference, July 2001). In *Algorithms for Approximation IV* (edited by I. J. Anderson, J. Levesley and J. C. Mason), 138–145 (University of Huddersfield, UK, July 2002).
7. F.Pavese, P.Ciarlini, Classes of inter-comparisons and the case of temperature standards, *PTB Berichte*, PTB-IT-10 (2003), 63-76.
8. F.Pavese, *Final Report of the International Intercomparison of fixed points by means of sealed cells: 13.81 K to 90.686 K*, Monograph 84/4, Bureau International des Poids et Mesures, Sèvres, 1984, 219.
9. CIPM, *Mutual Recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes*, Bureau International des Poids et Mesures, Sèvres, 1999.
10. Final Report of Key Comparisons CCT-K2, Summary: Steele A. G., Fellmuth B., Head D. I., Hermier Y., Kang K. H., Steur P. P. M., Tew W. L., *Metrologia*, 2002, 39, 551–572.
11. Final Report of Key Comparisons CCT-K4, Summary: H. G. Nubbemeyer and J. Fischer, *Metrologia*, 2002, 39, Techn. Suppl. 03001.
12. A.G.Steele, K.D. Hill, R.J. Douglas, Pooled data distributions: graphical and statistical tools for examining comparison reference values, *This Book*.

HOMOTOPIC SOLUTION OF EW-TLS PROBLEMS*

MARIA LUISA RASTELLO

*Istituto Elettrotecnico Nazionale Galileo Ferraris,
Strada delle Cacce 91, I-10135 Torino, Italy. E-mail: email rastello@ien.it*

AMEDEO PREMOLI

*Dipartimento di Elettronica e Informazione, Politecnico di Milano,
Piazza Leonardo Da Vinci 32, I-20133 Milano, Italy.*

This paper presents a new homotopic algorithm for solving Elementwise-Weighted Total-Least-Squares (EW-TLS) problems. For this class of problems the assumption of identical variances of data errors, typical of classical TLS problems, is removed, but the solution is not available in a closed form. The proposed iterative algorithm minimizes an *ad hoc* parametric weighted Frobenius norm of errors. The gradual increase of the continuation parameter from 0 to 1 allows one to overcome the crucial choice of the starting point. Some numerical examples show the capabilities of this algorithm in solving EW-TLS problems.

1. Introduction

In different scientific areas, various problems are formulated as overdetermined systems of linear equations of the type

$$A x \approx b, \quad (1)$$

where the entries of $A \in \mathbb{R}^{m \times n}$, with $n < m$, and $b \in \mathbb{R}^m$ are the noisy data of the problem and $x \in \mathbb{R}^n$ is the unknown vector; hereafter, set $\mathcal{I} \equiv \{1, 2, \dots, m\}$ groups the indices of scalar equations embedded in Eq. (1) and set $\mathcal{J} \equiv \{1, 2, \dots, n\}$ groups the indices of unknowns.

In the *Ordinary Least Squares* (OLS) problems, independently distributed (i.d.) errors are assumed to affect only vector b . Methods for estimating the effect of this assumption on the OLS solution are given in ¹. The *Total Least Squares* (TLS) problems, introduced in ², take into account i.d. data errors in both A and b , under the assumption of identical variances

*Work partially funded under EU SofTools_MetroNet Contract N. G6RT-CT-2001-05061

(i.i.d. errors). Their solutions have been found in a closed form, by using the *Singular Value Decomposition* (SVD), as pointed out in ³, thus avoiding the use of an iterative algorithm. A full review on general TLS problems can be found in ⁴. The case opposite, but less popular, to the OLS problem is the *Data Least Squares* (DLS) problem ⁵, where i.i.d. errors affect only the $m \times n$ entries of A .

Unfortunately, in many problems A and b represent various physical quantities and, therefore, are measured with different accuracy. As a consequence, entries a_{ij} of A and b_i of b are affected by independently distributed (i.d.) errors, Δa_{ij} and Δb_i , with zero averages and different variances $\sigma_{ij}^2 \forall ij \in \mathcal{I} \times \mathcal{J}$ and $\sigma_{ib}^2 \forall i \in \mathcal{I}$. Hereafter, the corresponding standard deviations are cast in the following matrix and vector

$$\Sigma \equiv \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \dots & \sigma_{mn} \end{bmatrix} \quad \text{and} \quad \sigma_b \equiv \begin{bmatrix} \sigma_{1b} \\ \sigma_{2b} \\ \vdots \\ \sigma_{mb} \end{bmatrix},$$

respectively. These problems are called *Elementwise-Weighted Total-Least-Squares* (EW-TLS) problems ⁶ and result to be of wide interest in many measurement applications, as evidenced in ⁷ and ⁸.

A solution in a closed form is not known for generic EW-TLS problems. This paper proposes an original – to authors' knowledge – and robust iterative algorithm solving these problems. It is based on the minimization of the weighted Frobenius norm (WFN) of errors ⁹, that results to be equal to the sum of ratios of suitable quadratic forms versus x . This objective function is convex constrained to a finite subregion including the expected minimum, and requires an iterative algorithm to find the solution. However, the choice of a good starting point is critical, because the boundary of the convexity subregion cannot be easily determined.

In the here-proposed algorithm, the original WFN is preliminarily expanded into a parametric family of functions by introducing a suitable continuation parameter $\eta \in [0, 1]$ and approximated by the truncated second-order Taylor formula. Initially the minimum corresponding to $\eta = 0$ is determined in a closed form as the problem reduces to the OWLS one and, then, exploited as the starting point to determine the minimum of the objective function for η close enough to 0. Then, according to a homotopic strategy, η is gradually increased up to $\eta = 1$, to find the minimum of the original WFN. In other words, the algorithm follows, step by step, the movement of the minimum versus η in order to center the solution of the

original EW-TLS problem. At the k -th step the minimum of the objective function is determined by few Newton (algorithm) sub-steps, based on its truncated Taylor formula in the neighborhood of the solution x^* of the previous substep.

When a non-definite positive Hessian matrix is met in the Taylor formula, the iterations of Newton algorithm are stopped and the algorithm is restarted with a lower value of η to keep the current solution inside the convexity region. In this way, the critical choice of the starting point is overcome.

Numerical results show the capabilities, in particular the robustness and the local convergence, of the algorithm. Unfortunately, a proof showing the global convergence is not available. However, some numerical tests of the algorithm on specific EW-TLS problems, whose global minimum is known, show the coincidence of the solution with the global minimum.

2. Problem statement

As in ⁹, we introduce the WFN of data errors on A and B

$$F(x, \Delta A, \Delta b) = \sum_{i \in \mathcal{I}} (\Delta a_i^T S_i^{-1} \Delta a_i + \Delta b_i^2 \sigma_{ib}^{-2}) , \quad (2)$$

where $\Delta a_i = [\Delta a_{i1} \ \Delta a_{i2} \ \dots \ \Delta a_{in}]^T$, $\Delta b = [\Delta b_1 \ \Delta b_2 \ \dots \ \Delta b_m]^T$ and $S_i \equiv \text{diag}(\sigma_{i1}^2, \sigma_{i2}^2, \dots, \sigma_{in}^2)$ ($\forall i \in \mathcal{I}$). Solving the EW-TLS problem consists in finding the optimal values \tilde{x} , $\Delta \tilde{A}$ and $\Delta \tilde{b}$ of x , ΔA and Δb , respectively, by minimizing $F(x, \Delta A, \Delta b)$

$$\min_{\langle x, \Delta A, \Delta b \rangle} F(x, \Delta A, \Delta b) \quad \text{subject to} \quad (A - \Delta A) x = b - \Delta b . \quad (3)$$

Note that \tilde{x} , $\Delta \tilde{A}$ and $\Delta \tilde{b}$ are invariant with respect to any arbitrary scaling factor in the variances. When $\sigma_{ij} = 0$, the corresponding error Δa_{ij} is excluded from the WFN in Eq. (2). If $S_i = 0$ and $\sigma_{ib}^2 = 0$, the i -th scalar equation in Eq. (1) becomes exact and “ \approx ” must be replaced by “ $=$ ”. The possible equations with $S_i = 0$ and $\sigma_{ib}^2 = 0$ play the role of “exact” constraints on x . In principle, these cases can be treated as limit cases of the EW-TLS problems, but their solution requires a different formulation, and, therefore, their investigation is out of the aim of the present paper. By choosing specific structures of Σ and σ_b the above formulation of EW-TLS problems includes :

- (1) Ordinary Least Squares (OLS) problems : when $\sigma_{ij} = 0 \ \forall i \in \mathcal{I} \times \mathcal{J}$ and σ_{ib} are nonzero and invariant $\forall i \in \mathcal{I}$.

- (2) Ordinary Weighted Least Squares (OWLS) problems : when $\sigma_{ij} = 0 \forall ij \in \mathcal{I} \times \mathcal{J}$ and σ_{ib} are nonzero and depending on $i \in \mathcal{I}$.
- (3) Total Least Squares (TLS) problems : when $\sigma_{ij} \forall ij \in \mathcal{I} \times \mathcal{J}$ and $\sigma_{ib} \forall i \in \mathcal{I}$ are nonzero and coincident.
- (4) Data Least Squares (DLS) problems : when $\sigma_{ij} \forall ij \in \mathcal{I} \times \mathcal{J}$ are nonzero and coincident and $\sigma_{ib} = 0 \forall i \in \mathcal{I}$.
- (5) Relative Error Total Least Squares (RE-TLS) problems ⁶ : when $\sigma_{ij} \propto |a_{ij}| \forall ij \in \mathcal{I} \times \mathcal{J}$ and $\sigma_{ib} \propto |b_i| \forall i \in \mathcal{I}$.
- (6) Weighted Total Least Squares (WTLS) problems ^{2 10} : when the variances verify the relations $\sigma_{ij}^2 = \sigma_{ib} \sigma_j' \forall ij \in \mathcal{I} \times \mathcal{J}$.

On the other hand, EW-TLS problems can be regarded as particular Structured-TLS problems ^{6 11}. However, in the latter the number of uncertain free parameters is upper constrained by the number of the $m \times (n+1)$ elements of system (1) and is, in practice, much fewer than in a generic EW-TLS problem.

With reference to Eq. (3), we substitute $\Delta b = \Delta A x - A x + b$ into $F(x, \Delta A, \Delta b)$ and obtain a new expression of the WFN

$$F(x, \Delta A) = \sum_{i \in \mathcal{I}} \left[\Delta a_i^T S_i^{-1} \Delta a_i + (x^T \Delta a_i - x^T a_i + b_i)^2 \sigma_{ib}^{-2} \right] . \quad (4)$$

To minimize the above WFN, we determine, for a given value of x , the m gradients of $F(x, \Delta A)$ with respect to the unknown n -vectors $\Delta a_1, \Delta a_2, \dots, \Delta a_m$

$$\frac{\partial F(x, \Delta A)}{\partial \Delta a_i} = 2 \left[S_i^{-1} \Delta a_i + (x^T \Delta a_i - x^T a_i + b_i) \sigma_{ib}^{-2} x \right] \quad \forall i \in \mathcal{I} .$$

Setting them to zero, we obtain the following m systems of n algebraic equations in the n -vectors $\Delta a_1, \Delta a_2, \dots, \Delta a_m$

$$[S_i^{-1} + x x^T \sigma_{ib}^{-2}] \Delta a_i = (x^T a_i - b_i) \sigma_{ib}^{-2} x \quad \forall i \in \mathcal{I} . \quad (5)$$

Any matrix in Eqs. (5) coincides with the sum of a diagonal matrix and a rank-one matrix, and its inverse is given by the closed form

$$[S_i^{-1} + x x^T \sigma_{ib}^{-2}]^{-1} = S_i - \frac{S_i x [S_i x]^T}{(x^T S_i x + \sigma_{ib}^2)} \quad i \in \mathcal{I} .$$

So, the solutions of systems in Eqs. (5) are

$$\Delta a_i = \left[\frac{(x^T a_i - b_i)}{(x^T S_i x + \sigma_{ib}^2)} \right] S_i x \quad i \in \mathcal{I} . \quad (6)$$

Hereafter, we assume that $x^T S_i x + \sigma_{ib}^2 > 0$, for any i and $\forall x$ in a neighborhood of the minimizer of $F(x, \Delta A)$. In practice, this limitation is not heavy.

By substituting the values Δa_i in Eqs. (6) into $F(x, \Delta A)$ (see Eq. 4), we obtain a new expression of the WFN, depending only on x

$$F(x) = \sum_{i \in \mathcal{I}} \left(\frac{x^T a_i - b_i}{x^T S_i x + \sigma_{ib}^2} \right)^2 (x^T S_i x + \sigma_{ib}^2) = \sum_{i \in \mathcal{I}} \frac{(x^T a_i - b_i)^2}{x^T S_i x + \sigma_{ib}^2} . \quad (7)$$

Then the minimization of $F(x, \Delta A)$ in Eq. (4) can be substituted by the less troublesome minimization of $F(x)$ with $x \in \mathbb{R}^n$

$$\min_{\langle x \rangle} F(x) . \quad (8)$$

Note that terms $(a_i^T x - b_i)^2$ and $(x^T S_i x + \sigma_{ib}^2)$, $i \in \mathcal{I}$, are non-negative and analytically continuous. If $\sigma_{ib} > 0$ the corresponding term $(x^T S_i x + \sigma_{ib}^2)$ is positive $\forall x \in \mathbb{R}^n$ and $F(x)$ is finite and analytically continuous $\forall x \in \mathbb{R}^n$. If $\sigma_{ib} = 0$ the corresponding term $(x^T S_i x + \sigma_{ib}^2)$ can be 0. For the values of x zeroing one of these terms, $F(x)$ is analytically discontinuous and tends to $+\infty$; for the remaining values of x , $F(x)$ is finite and analytically continuous.

Since problem in Eq. (8) cannot, in general, be solved in a closed form, an iterative (homotopic) algorithm is designed taking into account the peculiarities of the objective function. To obtain a fast convergent algorithm, we assume $F(x)$ to be strictly convex in the neighborhood $\mathcal{D}^* \subset \mathbb{R}^n$ of a given point x^* and approximate it by the 2nd-order truncated Taylor formula $\tilde{F}(x^* + \Delta x) \simeq F(x)$, with $\Delta x = x - x^*$, $\forall x \in \mathcal{D}^*$. Under the convexity assumption, a stationary point of $F(x)$ (defined by the nullity of the gradient) is a local minimum too. This assumption implies the positive definiteness of the Hessian matrix.

Since $F(x)$ in Eq. (7) is not, in general, convex $\forall x \in \mathbb{R}^n$, the choice of the starting point x^* is crucial to obtain an efficient algorithm. However, there exists a convexity region $\mathcal{R}_x \subset \mathbb{R}^n$, including the minimizer \tilde{x} , over which $F(x)$ is strictly convex. Unfortunately, the boundary of \mathcal{R}_x is not easy to be determined.

To overcome this drawback, we introduce a continuation parameter η and expand $F(x)$ in Eq. (7) into a parametric family $\bar{F}(x, \eta)$, called Parametric Weighted Frobenius Norm (PWFN)

$$\bar{F}(x, \eta) = \sum_{i \in \mathcal{I}} \frac{(x^T a_i - b_i)^2}{\eta x^T S_i x + \sigma_{ib}^2} = \sum_{i=1}^m \frac{C_i(x)}{D_i(x, \eta)} \quad \text{with } \eta \in [0, 1] . \quad (9)$$

Since both the minimizer \tilde{x} and the convexity region \mathcal{R}_x vary with respect to η , in the sequel we will denote them by $\tilde{x}(\eta)$ and $\mathcal{R}_x(\eta)$, respectively. The curve described by $\tilde{x}(\eta)$ in \mathbb{R}^n , by varying η in $[0, 1]$, is called *minima locus*, while the region described by $\mathcal{R}_x(\eta)$ forms the *convexity funnel* in the \mathbb{R}^{n+1} space spanned by x and η , including the minima locus.

By examining $\bar{F}(x, \eta)$ in Eq. (9), we observe that $\bar{F}(x, 0)$ corresponds to the WFN of the OWLS problem derived from the original EW-TLS problem by assuming that all the entries of A are error-free, (i.e., $\sigma_{ij} = 0$, $\forall i, j \in \mathcal{I} \times \mathcal{J}$). In this case, $\tilde{x}(0)$ can be determined in a closed form and $\mathcal{R}_x(0)$ coincides with \mathbb{R}^n , because $\bar{F}(x, 0)$ reduces to a quadratic form in x .

By varying η from 0 to 1, $\bar{F}(x, \eta)$ represents the WFN of EW-TLS problems in which the variances of entries of A are proportional to those of the original EW-TLS problem in Eq. (2), i.e. $\eta \sigma_{ij}^2 \forall i, j \in \mathcal{I} \times \mathcal{J}$. For $\eta = 1$, $\bar{F}(x, 1)$ coincides with $F(x)$ in Eq. (7).

Moreover, the convexity region $\mathcal{R}_x(\eta)$ contracts from $\mathcal{R}_x(0) \equiv \mathbb{R}^n$ to a finite region in $\mathcal{R}_x(1)$. If $\tilde{x}(\eta^*) \in \mathcal{R}_x(\eta^*)$, for the sake of analytical continuity, we have $\tilde{x}(\eta^* + \Delta\eta) \in \mathcal{R}_x(\eta^*)$ for $\Delta\eta$ small enough.

Note that, while the points on the surface delimiting the funnel are characterized by the positive semidefiniteness of the Hessian, the points within the funnel are characterized by the positive definiteness of the Hessian and the Hessian matrix can be decomposed into the product $U^T U$, where U is an upper triangular matrix. The possible non-existence of factorization $U^T U$ implies that the matrix is neither positive definite nor positive semidefinite; thus the current solution x is external to the funnel.

3. The convexity funnel algorithm

In this section we derive the Newton algorithm to minimize the PWFN $\bar{F}(x, \eta)$, for a given η . To this aim we introduce the following notations: $\nabla_x \bar{F}(x, \eta)$ and $\nabla_x^T \bar{F}(x, \eta)$ are the column and row gradient vectors of $\bar{F}(x, \eta)$ versus x , while $\nabla_x \nabla_x^T \bar{F}(x, \eta)$ is the Hessian matrix of $\bar{F}(x, \eta)$ versus x . Note that in these definitions the continuation parameter η is kept invariant. The second-order truncated Taylor formula $\tilde{F}(x^* + \Delta x, \eta)$ approximating $\bar{F}(x, \eta)$ in the neighborhood of a generic point x^* is given by

$$\tilde{F}(x^* + \Delta x, \eta) = \bar{F}(x^*, \eta) + \nabla_x^T \bar{F}(x^*, \eta) \Delta x + 1/2 \Delta x^T \nabla_x \nabla_x^T \bar{F}(x^*, \eta) \Delta x, \quad (10)$$

where

$$\nabla_x^T \bar{F}(x^*, \eta) = \sum_{i=1}^m \frac{2(a_i^T x^* - b_i)}{\eta x^{*T} S_i x^* + \sigma_{ib}^2} a_i^T - \sum_{i=1}^m \frac{2\eta (a_i^T x^* - b_i)^2}{(\eta x^{*T} S_i x^* + \sigma_{ib}^2)^2} x^{*T} S_i, \quad (11)$$

and

$$\begin{aligned} \nabla_x \nabla_x^T \bar{F}(x^*, \eta) &= \sum_{i=1}^m \frac{2}{\eta x^{*T} S_i x^* + \sigma_{ib}^2} a_i a_i^T - \sum_{i=1}^m \frac{2\eta (a_i^T x^* - b_i)^2}{(\eta x^{*T} S_i x^* + \sigma_{ib}^2)^2} S_i \\ &\quad - \sum_{i=1}^m \frac{4\eta (a_i^T x^* - b_i)}{(\eta x^{*T} S_i x^* + \sigma_{ib}^2)^2} (S_i x^* a_i^T + a_i x^{*T} S_i) \\ &\quad + \sum_{i=1}^m \frac{8\eta^2 (a_i^T x^* - b_i)^2}{(\eta x^{*T} S_i x^* + \sigma_{ib}^2)^3} S_i x^* x^{*T} S_i. \end{aligned} \quad (12)$$

The second-order Taylor formula in Eq. (10) is exploited to realize a pure (quadratically convergent) iterative Newton algorithm. If x^* is a point close enough to the searched minimum, each single Newton sub-step consists in determining an increment $\Delta x = x - x^*$ under the assumption that the Taylor expansion be good enough and that the objective function be strictly convex in the neighborhood of x^* .

$$\nabla_{\Delta x} \tilde{F}(x^* + \Delta x, \eta) = \nabla_x \bar{F}(x^*, \eta) + \nabla_x \nabla_x^T \bar{F}(x^*, \eta) \Delta x. \quad (13)$$

Each stationary point is given then by

$$\Delta x = - [\nabla_x \nabla_x^T \bar{F}(x^*, \eta)]^{-1} \nabla_x \bar{F}(x^*, \eta). \quad (14)$$

If $\bar{F}(x, \eta)$ results to be convex at $x = x^* + \Delta x$, we set $x^* = x^* + \Delta x$ and execute a further Newton substep. Otherwise, we stop the Newton algorithm and restart it, by choosing a lower value for η . This trick is fundamental in order to maintain x^* within the convexity funnel. More precisely, let η^* be the highest value of η for which we know the corresponding minimum $x^{**} \equiv \tilde{x}(\eta^*)$ and $\eta_a > \eta^*$ be the value of η for which the Newton algorithm, starting from x^{**} , does not converge. We repeat the Newton algorithm, starting again from x^{**} , but choosing a lower value $\eta_b \in (\eta^*, \eta_a)$. This decrease of η makes easier the convergence of the Newton algorithm because the searched minimum $\tilde{x}(\eta_b)$ is closer to x^{**} than the previous $\tilde{x}(\eta_a)$. If the Newton algorithm does not converge also for $\eta = \eta_b$, η_b is further decreased until a value $\eta_b > \eta^*$ assuring the convergence is obtained. The success of this homotopic procedure is assured by the analytical continuity of $\tilde{x}(\eta)$,

due to the analytical continuity of $\bar{F}(x, \eta)$ for $\forall x$ near to its minima. So, we can set $\eta^* = \eta_b$ and $\eta_a = \eta^* + \Delta\eta$ and repeat the procedure with a higher values of η^* .

The choice of $\Delta\eta$ is crucial : for $\Delta\eta$ small enough, the convergence of the Newton algorithm is assured, but many increments of η can be necessary to reach the final value $\eta = 1$. On the other hand, for $\Delta\eta$ large enough, few increments of η can be sufficient in principle, but the Newton algorithm cannot converge. To realize a good trade-off, we choose the starting point of a Newton algorithm for a given η according to some extrapolation techniques.

If we know a set of minimizers $\tilde{x}(\eta_\nu), \tilde{x}(\eta_{\nu+1}), \dots, \tilde{x}(\eta_\mu)$ corresponding to the set of values $0 < \eta_\nu < \eta_{\nu+1} < \dots < \eta_\mu$, thus, we can estimate the value of $\tilde{x}(\eta_{\mu+1})$ corresponding to an assigned value $\eta_{\mu+1} > \eta_\mu$ by extrapolating the known values $\tilde{x}(\eta_\nu), \tilde{x}(\eta_{\nu+1}), \dots, \tilde{x}(\eta_\mu)$. The extrapolation is obtained by adopting n simple functions, one for each entry of x . These extrapolation formulas allow one to choose the starting point for determining the minimizer $\tilde{x}(\eta_{\mu+1})$ so accurately that the Newton method supplies the true value in a few iterations.

The initial value of η can be chosen on the basis of experience. In the next section we determine it by an empirical formula. However, this value is not at all critical because it is varied by the algorithm itself when the Newton algorithm does not converge.

The convergence of the Newton algorithm is controlled by distance $\|\Delta x\|$: for $\|\Delta x\| < \epsilon^2 \|x^*\|$ we ascertain that x^* is an equilibrium point of $\bar{F}(x, \eta)$ and the positive definiteness of Hessian matrix $\nabla_x \nabla_x^T \bar{F}(x^*, \eta) = U^T U$ assures that x^* is inside the convexity funnel and, hence, it is a minimum of $\bar{F}(x, \eta)$. In the algorithm we introduce the desired relative accuracy ϵ_0 of partial minima (i.e, for $\eta < 1$), and the accuracy $\epsilon_1 < \epsilon_0$ of the final minimum (i.e, for $\eta = 1$).

If the distance $\|\Delta x^*\|$ between the minima related to the k_η -th value of η and the $(k_\eta - 1)$ -th one is small enough, the next value of η is set to 1. Otherwise, some strategy has to be determined in order to improve the speed of the algorithm. Let ϵ_2 be the accuracy of the distance to distinguish between two successive minima, k_η is the counter of the increments of η and k_x is the counter of the substeps within a single Newton algorithm.

The main computational effort is required in calculating and factorizing the Hessian matrix. Several multiplications and additions of matrices and vectors are involved. However, the matrices in products are rank-one (matrices $a_i a_i^T, x x^T, a_i x^T, x a_i^T$ in Eq. (12)) or diagonal (matrices

S_1, S_2, \dots, S_m in Eq. (12)), due to the structure of the objective function $\bar{F}(x, \eta)$ in Eq. (9). Generic non-sparse matrices are involved only in less expensive summations versus ι . This fact makes the calculation of the Hessian less CPU expensive than that of a generic $\Re^n \rightarrow \Re$ function. The LU factorization of the Hessian matrix is realized by the standard Cholesky algorithm.

4. Numerical results

All examples are concerning with specifically noisy matrix A and vector b derived from a unique error-free 10×2 matrix and error-free vector, by adding random errors to all entries. The initial value of $\Delta\eta$ starting the algorithm is determined by the empirical formula

$$\Delta\eta = \min_{\langle i \rangle} \frac{\sigma_{i,n+1}^2}{\sum_{j=1}^n \sigma_{i,j}^2} .$$

In all examples the accuracy required for stopping the intermediate iterative Newton algorithm is $\epsilon_0 = 0.05$, while the accuracy for stopping the overall algorithm is $\epsilon_1 = 10^{-12}$; $\epsilon_2 = 0.05$ is the accuracy for the distance between minima related to two successive increments of η .

The examples are characterized by the same standard deviations of the respective errors

$$\Sigma = \begin{bmatrix} 4.0000 & 0.1000 \\ 3.5667 & 0.5333 \\ 3.1333 & 0.9667 \\ 2.7000 & 1.4000 \\ 2.2667 & 1.8333 \\ 1.8333 & 2.2667 \\ 1.4000 & 2.7000 \\ 0.9667 & 3.1333 \\ 0.5333 & 3.5667 \\ 0.1000 & 4.0000 \end{bmatrix} \quad \text{and} \quad \sigma_b = \begin{bmatrix} 0.0100 \\ 0.0100 \\ 0.0100 \\ 0.0100 \\ 0.0100 \\ 0.0100 \\ 0.0100 \\ 0.0100 \\ 0.0100 \\ 0.0100 \end{bmatrix} ,$$

whereas data matrix A and data vector b are different. This induces the same starting value $\Delta\eta = 0.00000625$ but the behaviour of the algorithm results to be rather different.

For

$$A = \begin{bmatrix} 2.59682756412627 & 9.14729111791493 \\ 6.82215678878727 & 9.19327278015292 \\ 3.46230227137220 & 2.74926114730709 \\ 3.20353427391860 & 1.66816071164268 \\ 10.89991665143950 & -1.61023973136850 \\ -0.37203854346990 & 6.13279826630644 \\ 4.65527261492859 & 4.55510312153970 \\ 0.65763908669889 & 10.76259377967540 \\ 3.06849059794170 & 7.85144473997031 \\ 2.97487063693829 & 13.19952354957530 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 67.008389261 \\ 63.003350575 \\ 26.995457326 \\ 2.997805620 \\ 9.001552853 \\ 42.990951419 \\ 51.989789658 \\ 56.995640572 \\ 46.013382142 \\ 24.010696135 \end{bmatrix},$$

the algorithm converges in the predicted way as shown in Table 1. Table 1 shows the numerical results illustrating the behavior and performances of the algorithm. In particular Table 1 shows the starting point x and the obtained point x^* for each Newton substep. The last column reports the squared modulus of the difference between the minima related to two successive increments of η .

TABLE 1

	(k_η, k_x)	η	x_1^*	x_2^*	$\ \Delta x^*\ $
1		OWLS	2.108078099	4.539951373	-
2	(1,1)	0.00000625	1.772811024	5.878032581	-
3	(1,2)		1.621103047	6.261642009	-
4	(1,3)		1.659813180	6.233089613	0.124E+01
5	(2,1)	0.00002498	1.771939925	6.193504350	-
6	(2,2)		1.774217749	6.195734238	0.851E-01
7	(3,1)	0.00008120	1.806730504	6.183439118	-
8	(3,2)		1.806938644	6.183581201	0.247E-01
9	(4,1)	1.00000000	1.822081254	6.177798235	-
10	(4,2)		1.822130963	6.177825737	-
11	(4,3)		1.822130968	6.177825736	-
12	(4,4)		1.822130968	6.177825736	0.115E-01

The initial value of $\Delta\eta$ is small enough and four increments of η are

required. On the contrary, for

$$A = \begin{bmatrix} 4.67143070625183 & 8.90028320737129 \\ 12.49822068903150 & 8.27703510323791 \\ 6.93862923994614 & 3.61443195076070 \\ 3.40768477129794 & 0.14275866741392 \\ 8.80844351480653 & 1.01201437135240 \\ 3.73390581806528 & 2.50419998991749 \\ 2.27239857108500 & 5.85511824073525 \\ 1.30501094112949 & 4.87872627074283 \\ 4.68577780114022 & 3.45851362432531 \\ 2.85242249761820 & -3.69955579089445 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 67.002621315 \\ 62.997016898 \\ 27.006431377 \\ 2.982733880 \\ 8.991048680 \\ 43.008023017 \\ 52.004382496 \\ 56.987487133 \\ 46.007503142 \\ 24.026246284 \end{bmatrix},$$

the initial value of $\Delta\eta$ results to be too large. In fact, not only for $\eta = 0.00000625$ but also for $\eta = 0.00000125$ the convergence is not reached because x exits the funnel, as shown in Table 2. However, the convergence of the Newton substeps is obtained for $\eta = 0.00000025$. Thereafter, seven increments of η are required to reach the final minimum.

TABLE 2

	(k_η, k_x)	η	x_1^*	x_2^*	$\ \Delta x^*\ $
1	OWLS		1.769743953	6.378399634	-
2	(1,1)	0.00000625	-28.572823559	23.451717413	Newton algorithm stopped
3	(1,2)				
4	(1,1)		-1.626309116	8.889327409	-
5	(1,2)	0.00000125	5.577592673	3.277287639	-
6	(1,3)				Newton algorithm stopped
7	(1,1)		-0.126649107	8.397578230	
8	(1,2)	0.00000025	-0.062206462	8.450164555	-
9	(1,3)		-0.062144879	8.450585216	0.196E+01
10	(2,1)	0.00000100	-0.242721491	8.257340967	-
11	(2,2)		-0.227987309	8.251298262	0.183E+00
12	(3,1)	0.00000325	-0.015263290	7.784231228	-
13	(3,2)		-0.112771278	7.890215060	-
14	(3,3)		-0.118013663	7.895949231	0.263E+00
15	(4,1)	0.00000999	0.134357432	7.532394610	-
16	(4,2)		0.002995922	7.654061812	-
17	(4,3)		-0.002644704	7.659147132	0.186E+00
18	(5,1)	0.00003023	0.091304845	7.526564659	-
19	(5,2)		0.057380884	7.550772259	0.876E-01
20	(6,1)	0.00009094	0.080115942	7.511082640	-
21	(6,2)		0.080080201	7.510709438	0.326E-01
22	(7,1)	1.00000000	0.083490152	7.494310985	-
23	(7,2)		0.092464231	7.489498343	-
24	(7,3)		0.092426145	7.489521438	-
25	(7,4)		0.092426145	7.489521437	-
26	(7,5)		0.092426145	7.489521437	0.173E-01

5. Conclusions

The here-proposed convexity funnel algorithm, implementing a homotopic strategy, solves iteratively general Total-Least-Squares problems, in which errors are assumed to be elementwise independently distributed but with different variances (the so-called Elementwise-Weighted Total-Least-Squares problems). It is based on the introduction of a parametric weighted Frobenius norm of the errors. The crucial choice of the starting point for the minimization of this norm is overcome by gradually increasing the continuation parameter appearing in the norm. All examples prove that the algorithm is locally convergent and, in the special case of TLS setup, the EW-TLS estimator coincides with the corresponding TLS estimator.

In this paper, no rigorous proof assures that the solution supplied by the funnel algorithm is the global minimum of the weighted Frobenius norm

of errors. However, the confinement of the current point within the funnel, during the whole procedure, suggests the coincidence of the solution with the global minimum.

Some tests on realistic problems show the capabilities of the algorithm.

References

1. S.D. Hodges and P.G. Moore. Data uncertainties and least square regression. *Applied Statistics*, (21): 185-195, 1972.
2. G.H. Golub and C.F. Van Loan. An analysis of the total least squares problem. *SIAM Journal of Numerical Analysis*: 883-893, 1980.
3. G.H. Golub and C. Reinsch. Singular value decomposition and least squares solutions. *Nu. Math.*, (14): 403-420, 1970.
4. S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: Computational Aspects and Analysis*. Frontiers in Applied Mathematics. SIAM Philadelphia, 1991.
5. R.D. Degroat and E. Dowling. The data least squares problem and channel equalization. *IEEE Trans. on Signal Processing*, 41(1): 407-411, 1993.
6. B. De Moor. Structured total least squares and L_2 approximation problems. *Linear Algebra and its Applications*, 188, 189: 163-205, 1993.
7. A. Premoli, M.L. Rastello and G. Cirrincione. A new approach to total least-squares techniques for metrological applications. In P. Ciarlini, M.G. Cox, F. Pavese and D. Richter, editors, *Advanced Mathematical Tools in Metrology II, Series on Advances in Mathematics for Applied Sciences Vol. 40*: 206-215, World Scientific, 1996.
8. G. Cirrincione, S. Van Huffel, A. Premoli and M.L. Rastello. Iteratively reweighted total least squares algorithms for different variances in observations and data. In P. Ciarlini, M.G. Cox, E. Felipe, F. Pavese and D. Richter, editors, *Advanced Mathematical Tools in Metrology V, Series on Advances in Mathematics for Applied Sciences Vol. 57*: 77-84, World Scientific, 2001.
9. A. Premoli and M. L. Rastello. The parametric quadratic form method for solving TLS problems with elementwise weighting. in "Total Least Squares and Errors-in-Variables Modeling, S. Van Huffel and P. Lemmerling Editors, 67-76, Kluwer Academics Publ. 2002.
10. B.D. Rao. Unified treatment of LS, TLS and truncated SVD methods using a weighted TLS framework. In S. Van Huffel, editor, *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling*: 11-20, SIAM, 1997.
11. P. Lemmerling, S. Van Huffel and B. De Moor. Structured total least squares problems: formulation, algorithm and applications. In S. Van Huffel, editor, *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling*: 215-223, SIAM, 1997.

POOLED DATA DISTRIBUTIONS: GRAPHICAL AND STATISTICAL TOOLS FOR EXAMINING COMPARISON REFERENCE VALUES

A.G. STEELE, K.D. HILL, R.J. DOUGLAS

*National Research Council of Canada
1200 Montreal Road, Ottawa ON K1A 0R6, Canada
E-mail: alan.steele@nrc.ca*

Measurement comparison data sets are generally summarized using a simple statistical reference value calculated from the pool of the participants' results. This reference value can become the standard against which the performance of the participating laboratories is judged. Consideration of the comparison data sets, particularly with regard to the consequences and implications of such data pooling, can allow informed decisions regarding the appropriateness of choosing a simple statistical reference value. Recent key comparison results drawn from the BIPM database are examined to illustrate the nature of the problem, and the utility of a simple approach to creating pooled data distributions. We show how to use detailed analysis when arguing in favor of a KCRV, or when deciding that a KCRV is not always warranted for the particular data sets obtained experimentally.

1. Introduction

Among the many tasks facing international metrology today is the conduct and analysis of key comparisons. The goal is to quantify the degrees of equivalence among realizations of the system of units and the primary measurement techniques in each field.^a These comparisons are summarized in published tables that relate each participating laboratory result to a Key Comparison Reference Value (KCRV) that is, in general, an aggregate statistical estimator such as the mean (weighted uniformly or otherwise) or the median of the quoted laboratory results (both the values and the standard deviations).^b

We present a simple technique for displaying and interpreting the population of results, and for evaluating different statistical approaches to determining when and when not to use a KCRV that is calculated from the participants' data sets. This work amplifies the data pooling approach discussed in [1], and is similar to the mixture distributions explored in [2].

^a Available at <http://www.bipm.org/pdf/mra.pdf>

^b Available at <http://kcdb.bipm.org>

2. Data Pooling and Reference Values

In many statistical approaches to data analysis, the initial assumption is to consider that all of the participants' data represent individual samples from a single population. A coherent picture of the population mean and standard deviation can be built from the comparison data set that is fully consistent with the reported values and uncertainties. Most outlier-test protocols rely on this assumption to identify when and if a given laboratory result should be excluded, since its inclusion would violate this internal consistency.

Creating pooled data distributions tackles this problem from the opposite direction: rather than assuming a single (normal) distribution for the population, the independent distributions reported by each participant are summed directly, and the result is taken as representative of the underlying population of possible measurements reported by randomized participants.

When all participants' uncertainties are normal, it is straightforward to combine them and obtain expressions for the distributions of the population and for the simple and variance-weighted means. If finite degrees of freedom are provided, it can be more complex to combine the appropriate Student distributions analytically. The Welch-Satterthwaite approximation is often used to determine the effective degrees of freedom for a combination, for example.

Rigorous analysis in closed form is not always feasible. The standard uncertainty of the median, for example, has no closed form. Instead, the median absolute deviation (MAD), scaled by a multiplicative factor derived from normal statistics, is often taken as a working estimate of the standard deviation.

In general, a numerical approach to data pooling that is easy to use and fast to compute can provide metrologists with a useful visual and statistical tool to investigate many aspects of the comparison population and its estimators of central tendency as revealed in the reported measurements and uncertainties.

3. Monte Carlo Calculations

One of the most simple and versatile numerical methods for exploring these distributions and related statistics is the Monte Carlo technique, which relies on sampling to represent repeated fully randomized measurement comparisons. Each Monte Carlo "roll of the dice" for a comparison event involves randomly sampling every participant's reported distribution independently. In the language of bootstrap techniques, similar in implementation, this is "resampling without replacement". "Resampling with replacement" is inappropriate, since events would result with some repeat participants and some left out altogether.

3.1. Random Number Generation

Most texts on numerical analysis [3, 4] provide fast algorithms for generating uniformly distributed pseudorandom numbers. These generators provide streams with long or very long periodicity (in the range 2^{32} to 2^{363}), excellent uniformity and statistical properties, and can be seeded to allow repeated calculation using the same set of random numbers in different runs.

Transformation from such uniformly distributed random numbers to variates obeying any desired probability density function (PDF) can be effected using the corresponding cumulative distribution function (CDF). In this technique, the normalized CDF is computed for each participant from the reported mean values, standard deviations, and degrees of freedom for the appropriate PDF. Each uniformly distributed random variable may be transformed using a simple lookup procedure: take the uniform random number as the cumulative probability, and then the required variate may be read off of the abscissa directly or computed using linear interpolation between the bracketing pair of values. Figure 1 shows the both the PDF and CDF for a Student distribution with $\mu = -1$, $\sigma = 2$, and $\nu = 4$ as an aid to visualizing the transformation algorithm.

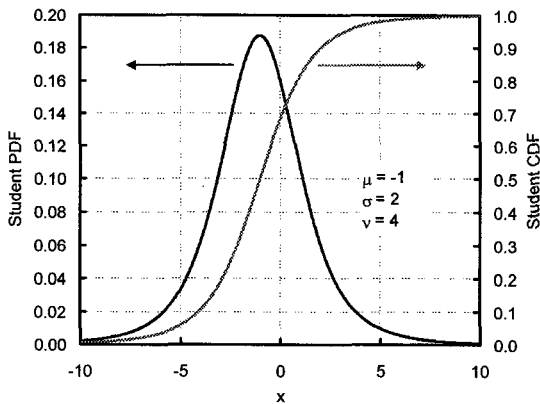


Figure 1. Plot of Student probability density function (dark line, left axis) and cumulative distribution function (light line, right axis). $CDF(x)$ is uniformly distributed, and may be used to transform uniform random numbers onto the Student (or any other) distribution.

We have found that calculating 10^4 points to describe each participant's CDF out to 7 or so standard deviations from their mean gives enough numerical resolution for evaluating comparison data. Far fewer points are required in the distribution histograms: sufficiently detailed curves can be obtained using 1000

bins. In this way, smooth graphs for the pooled distribution and the distributions of the simple mean, the variance weighted mean, the median, and many other simple estimators of central location can be generated using 10^6 or 10^7 events.

3.2. *Toolkit Implementation*

Our Excel Toolkit [5] has been extended to include macros that compute and display distributions.^c For a multi-laboratory comparison, measurements, uncertainties, degrees of freedom and inter-laboratory correlation coefficients are entered on an Excel worksheet. Toolkit macros make it simple to plot the individual participants' results, as well as the pooled distribution revealed in the comparison. Monte Carlo distributions of several common candidate reference values can also be calculated and plotted: the uniformly weighted (simple) mean, the inverse-variance weighted mean, and the median.

The Visual Basic for Applications (VBA) macros can be called directly from Excel, but most computation takes place in an external Dynamically Linked Library (DLL). These math-intensive functions have been written in FORTRAN, validated using a separate code base independently written in C. Using the toolkit, it is easy to generate and display pooled-data and reference-value distributions for a comparison with ten participants: ten million Monte Carlo resamplings of the comparison take only a few minutes on an ordinary PC.

4. A Worked Example

4.1. *Key Comparison Data Set*

As an example of the data pooling technique, consider the data in Table 1, extracted from Table 45 in the final report for the recent Key Comparison in Ultrasonic Vibration designated CCAUV.U-K1 in the BIPM database [6]. In this case, five laboratories performed an ultrasonic power measurement under reference conditions at 1.9 MHz. The laboratory results include their measured reference power (P_{Ref}), standard uncertainty (u), and degrees of freedom (ν). The pilot has reported that no correlations between participants have been identified, so that no covariances need to be taken into account. Although the final report summarizes the degrees of equivalence normalized by the reference value, here we have reverted to the usual format using the measurand's units.

^c Available at <http://inms-ienm.nrc-cnrc.gc.ca/qde>

These data are plotted in Figure 2. The Pilot determined that the data fail a consistency check that incorporates effective degrees of freedom, and therefore the median was chosen as the KCRV over the variance-weighted mean.

Table 1. Key Comparison data for five participants in CCAUV.U K1: ultrasonic power at 1.9 MHz, low power.

Lab	P_{Ref} (mW)	u (mW)	ν
PTB	97.4	0.84	8.3
NIST	99.0	0.64	6.3
NPL	97.6	1.01	11
CSIRO	114.5	6.75	6.7
NIM	94.0	1.16	12.5

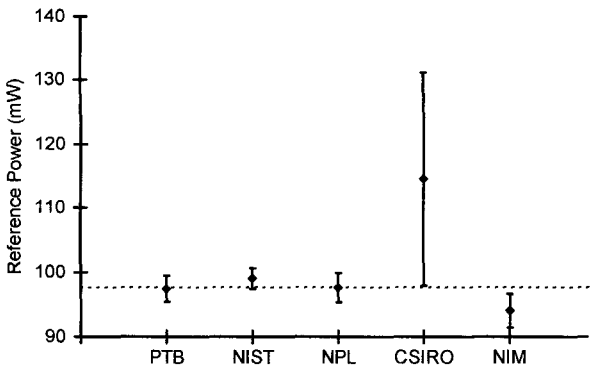


Figure 2. Plot of Table 1, with error bars shown at 95% level of confidence. The reference value, taken to be the median (97.6 mW), is indicated by the dotted line.

4.2. Monte Carlo Results

The pooled data distribution computed from 10^7 Monte Carlo events is shown in Figure 3, along with the distributions for the simple mean, the variance weighted mean, and the median. It is straightforward to extract the mean value and the 95% confidence interval for each of these distributions using the numerical results, as shown in Table 2.

Several insights into the comparison can be obtained from Figure 3. The pooled data distribution is multi-modal, and does not justify assuming a normal population for outlier testing. Furthermore, given the similar distributions, it is unclear how choosing the median over the weighted mean as the KCRV has any meaningful impact on the comparison summary. The standard deviation of the

median distribution is significantly smaller than the uncertainty calculated in the final report using the weighted MAD function: 0.74 mW rather than 1.32 mW.

Table 2. Candidate reference values for the data in Table 1.

Estimator	Distribution		95% Confidence Interval
	Mean	Std Dev	
Mean	100.50	1.40	[97.15, 103.80]
WtMean	97.75	0.42	[96.75, 98.70]
Median	97.60	0.74	[96.45, 99.30]

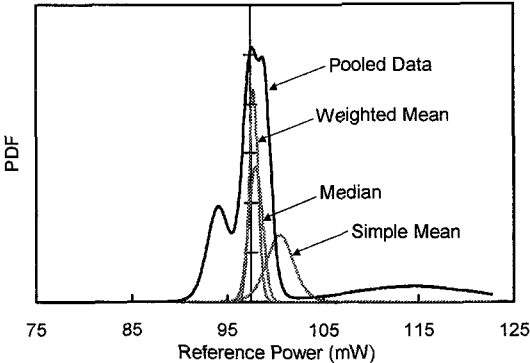


Figure 3. Plot of distributions derived from the data in Table 1. The pooled data is the sum of the individual distributions. The distributions for the simple mean, variance weighted mean, and median were calculated using Monte Carlo techniques. The KCRV, taken to be the instance median for this comparison, is indicated by the vertical axis.

5. Advantages of the Monte Carlo Approach

More important than its simplicity of use, the Monte Carlo technique offers full flexibility regarding methods to be studied.

With Monte Carlo methods, the statistical properties of any algorithmic method may be computed by writing a subroutine that implements the algorithm. Any such statistic on each of the participating laboratories (or pairs, or subgroups) may be generated to summarize their place in the comparison. Beyond treating the combined uncertainties as normal or Student distributions, and the correlations as normal distributions, individual terms in participants' uncertainty budgets, or the travel uncertainty of the artifact, may be assigned independent or shared random variates drawn from appropriate distributions.

5.1. *Covariances*

The covariance between a candidate reference value method and a participant laboratory can be found by Monte Carlo methods. After 10% of the events, the mean values are known, and the variance and covariance (away from the means) can be computed with the last 90% of the events. Similarly, the covariance of the weighted mean with the median might be calculated.

Where correlations between laboratories have been identified, the random numbers must be used with slightly greater care. For each simulated comparison, each laboratory value will be composed of its own unique independent part, and shared correlated parts. The combined uncertainty's Student distribution is decomposed into normally distributed random variables, each shared with one other Lab, and an independent Student-distributed random variable with a reduced degrees of freedom (such that the Welch-Satterthwaite approximation, applied to the combined uncertainty will return the appropriate Student distribution for the combined uncertainty). Some extra care is required to deal with negative correlation coefficients, adding the random variate to one Lab, and subtracting it from the other.

5.2. *Outlier Rejection Schemes*

In some comparisons, outlier rejection schemes are identified in the protocol; in others, this question is addressed during data analysis. Since the consequences for any participant laboratory of being identified as an outlier can be severe – including rejection of its calibration and measurement capability listing in Appendix C to the Mutual Recognition Arrangement – it is useful to explore these questions as carefully as possible.

Any algorithmic mechanism for outlier identification may be incorporated into the Monte Carlo simulation, and each laboratory may be assigned a probability of being so designated. Indeed, it is possible to track any comparable information in this fashion. For the example data set discussed in the previous Section, the probability of each participant being the median laboratory was found to be as follows: PTB (39%), NIST (17%), NPL (44%), CSIRO (0.5%), NIM (0.5%). It is interesting to note in particular that all participants, including the laboratory that failed the consistency check (NIM), have a non-zero chance of being the median laboratory.

5.3. *Other Candidate Reference Values*

The mean, weighted mean and median are three principal statistics used as estimators of central location, and considered as Key Comparison Reference

Values. At its last meeting, the CCT agreed to the use of a so-called “average reference value” (ARV) when determining which participants’ calibration and measurement capability claims related to CCT-K3 should receive further scrutiny prior to inclusion in Appendix C. In this particular context, the ARV is defined as the simple average of the usual three candidate reference values.

Figure 4 shows these distributions determined from the CCT-K3 data at the triple point of argon [7].^d The three usual estimators do not agree very well with each other, and do not provide a single reference value. The shape of the median distribution is interesting: it is very wide and flat: many participants have approximately equal chances of being a median lab. The ARV distribution is slightly skewed, a combination of the three distributions and their correlations.

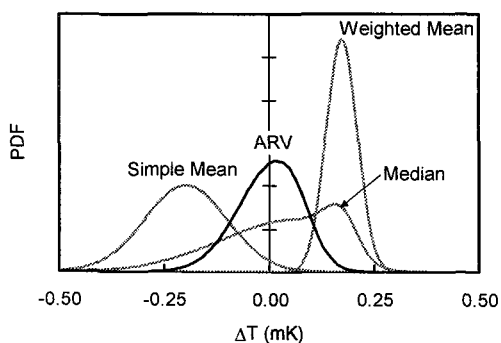


Figure 4. Plot of CCT-K3 distributions at the triple point of argon. The ARV (dark line) is the simple average of the mean, weighted mean, and median (light lines).

For CCT-K3, the technical experts agreed that no statistically meaningful KCRV could be assigned, since the pooled distribution revealed a complex underlying population with no simple estimator of central tendency. It is difficult to nominate a KCRV when different estimators each tend to favor or disfavor different participants. The simple linear combination to form the ARV is intended solely to serve as a single point of reference when identifying Appendix C claims that require further scrutiny. Such linear combinations of estimators of central tendency have been proposed and discussed elsewhere in the context of robust statistics [8]. The ease with which the Monte Carlo calculation yields the distribution for the ARV is illustrative of the power of this method, and may serve as encouragement to explore other, perhaps more complicated, estimators.

^d Available at <http://www.cstl.nist.gov/div836/836.05/papers/TN1450.pdf>

6. All-pairs Difference Distribution

Plotting the distributions for pooled data and estimators of central location enables visual insight into the comparison data set. To go beyond this qualitative level and quantitatively address the issue of when it is appropriate to nominate a KCRV calculated from the measurement results requires that the pair-difference distributions be calculated.

Figure 5 presents such graphs for the data summarized in Table 1. Each of the light curves represents the comparison from the perspective of an individual participant. The sum of these per-laboratory perspectives is shown as the dark curve, which is the all-pairs-difference pooled distribution, which is symmetric.

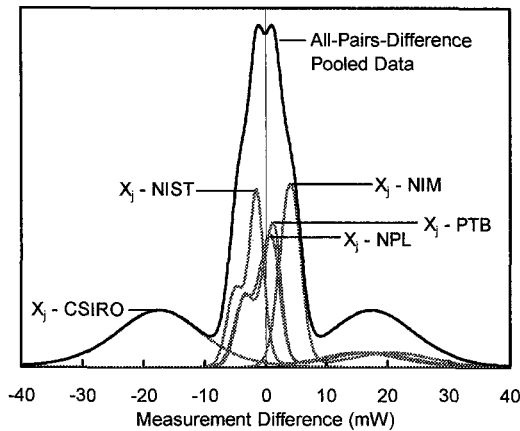


Figure 5. Pair difference distributions for the data in Table 1.

Similar to the exclusive statistics methods of [9], this pair-difference approach can be used to build an “exclusive chi-square” statistic for each participant by considering the appropriately scaled differences with respect to every other participant: $\chi_j^2 = (N-1)^{-1} \sum_{i=1, i \neq j}^N (x_i - x_j)^2 / (u_i^2 + u_j^2 - 2r_{ij}u_iu_j)$.

By averaging over all participants, we obtain a completely objective criterion that gives no special status to any laboratory’s results: the all-pairs variance expressed as a reduced chi-square: $\chi_r^2 = N^{-1} \sum_{j=1}^N \chi_j^2$.

This is the most general form of a chi-squared statistic for multi-participant comparisons. It is more powerful than the specific chi-squared test suggested in [10], which is tested against the weighted mean, since it is independent of *any* choice of KCRV: any constant shift applied to all laboratory values *exactly*

cancels out in this treatment. These quantities are shown in Table 3 for the data of Table 1. In each case, the probability of exceeding the computed value by chance is very small, indicating that the data set fails the chi-squared test, and that no single reference value can represent the underlying population.

Table 3. Reduced chi-squares from each Lab, and the all-pair-difference χ^2 - each with 4 degrees of freedom.

	PTB	NIST	NPL	CSIRO	NIM	APD
χ^2	3.57	5.78	3.25	6.65	8.57	5.57
$\Pr(\chi^2 > \chi^2_{\text{obs}})$	6.5×10^{-3}	1.2×10^{-4}	1.1×10^{-2}	2.4×10^{-5}	6.5×10^{-7}	1.8×10^{-4}

7. Conclusions

Monte Carlo methods have been exploited as a fast and convenient tool to create the distributions for any candidate reference value being considered as the representative statistic for summarizing a measurement comparison without the *a priori* assumption of data pooling in a single normal population.

References

1. A.G. Steele, K.D. Hill and R.J. Douglas, *Metrologia* **39**, 269 (2002).
2. F. Pavese and P. Ciarlini, *PTB Bericht Series IT: Proceedings of PTB Workshop on Uncertainty*, 63 (2002).
3. P.R. Bevington and D.K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 2nd Edition, Boston, USA, McGraw-Hill (1992).
4. D.E. Knuth, *The Art of Computer Programming Volume 2: Seminumerical Algorithms*, 3rd Edition, Reading, USA, Addison-Wesley (1997).
5. A.G. Steele, B.M. Wood and R.J. Douglas, *Advanced Mathematical and Computational Tools in Metrology* **5**, 320 (2001).
6. K. Beissner, *Metrologia Technical Supplement* **39**, 09001, (2002).
7. B. Mangum, G. Strouse, W. Guthrie, R. Pello, M. Stock, E. Renaot, Y. Hermier, G. Bonnier, P. Marcarino, K. Gam, K. Kang, Y. Kim, J. Nicholas, D. White, T. Dransfield, Y. Duan, Y. Qu, J. Connolly, R. Rusby, J. Gray, G. Sutton, D. Head, K. Hill, A. Steele, K. Nara, E. Tegeler, U. Noatsch, D. Heyer, B. Fellmuth, B. Thiele-Krivoj, S. Duris, A. Pokhodun, N. Moiseeva, A. Ivanova, M. de Groot, and J. Dubbeldam, *Metrologia* **39**, 179, (2002).
8. D.F. Andrews, P.J. Bickel, F.R. Hampel, P.J. Huber, W.H. Rogers and J.W. Tukey, *Robust Estimates of Location*, Princeton University Press, Princeton, USA, (1972).
9. A.G. Steele, B.M. Wood and R.J. Douglas, *Metrologia* **38**, 483, (2001).
10. M.G. Cox, *Metrologia* **39**, 589, (2002).

NUMERICAL UNCERTAINTY EVALUATION FOR COMPLEX-VALUED QUANTITIES: A CASE EXAMPLE*

LUCA CALLEGARO

Istituto Elettrotecnico Nazionale Galileo Ferraris (IEN)
Electrical Metrology Department
Strada delle Cacce, 91 - 10135 Torino, Italy
E-mail: lcallega@ien.it

F. PENNECCHI AND W. BICH

Istituto di Metrologia "Gustavo Colonnetti" (IMGC)
Mass Section, Strada delle Cacce, 73
10135 Torino, Italy
E-mail: f.pennecchi@imgc.cnr.it

The forthcoming Supplement 1 to the GUM: *Numerical methods for the propagation of distributions* proposes the use of Monte Carlo simulation for uncertainty evaluation. Here we apply a modified implementation of the proposed Monte Carlo Simulation to construct intervals of confidence for a complex-valued measurand. In particular, we analyze the so-called three-voltage method for impedance calibration, which relates complex-valued impedances to voltage moduli measurements. We compare and discuss the results obtained with those given by application of Bootstrap Resampling on the same model.

1. Introduction

Expressing physical quantities in the frequency domain requires dealing with variable values belonging to the field of complex numbers. A typical example is the phasor representation of voltages, currents and impedances in ac electrical measurements. The evaluation of measurement uncertainty with analytical methods is well established for real-valued quantities [1]. However, the application of these methods may be difficult when complex-valued quantities are involved in the measurement model. In this case, numerical techniques, such as Monte Carlo simulation (MCS) and bootstrap resampling (BR), may be advantageous since they can be implemented in a comparatively simple way, thus avoiding analytical difficulties.

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

As a case example, in this paper we study the impedance uncertainty evaluation when the three-voltage method of measurement [2] is used.

2. The three-voltage method

The three-voltage method is employed to measure an unknown impedance Z_X by comparison with a reference impedance Z_S . The principle is shown in Fig. 1: Z_X and Z_S are put in series and energized with a current I from a generator G . Voltages U_X and U_S develop on Z_X and Z_S ; vector linear combinations U_L of U_X and U_S can be generated with the aid of an inductive voltage divider IVD; in Fig. 1, the combinations $U = U_X + U_S$ and $U_M = (U_X - U_S)/2$ are shown. The rationale of the method lies in the much higher relative accuracy of rms (10^{-6}) versus vector (10^{-4}) voltmeters: by measuring three *rms* values (vector moduli) $|U_X|$, $|U_S|$, $|U_L|$, an equation $Z_X = f(|U_X|, |U_S|, |U_L|, Z_S)$ can be written to recover Z_X as a vector quantity.

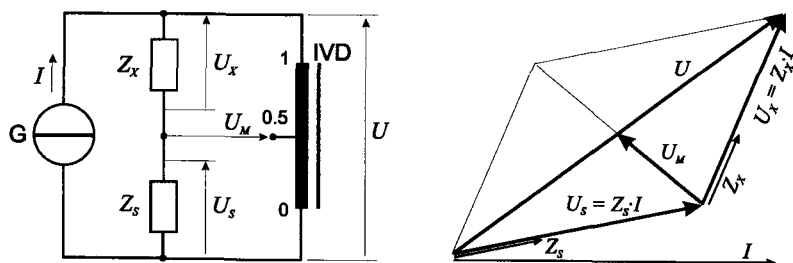


Figure 1. (left) Principle schematics of the three-voltage method. (right) Vector diagram.

When U_M is measured (a good choice for practical measurements), the model equation can be written as:

$$\alpha_X \equiv \frac{|U_X|}{|U_S|}; \alpha_M \equiv \frac{|U_M|}{|U_S|};$$

$$\alpha_A = \frac{1}{2} \left\{ \left[1 - (2\alpha_M - \alpha_X)^2 \right] \cdot \left[(2\alpha_M + \alpha_X)^2 - 1 \right] \right\}^{\frac{1}{2}}; \alpha_B = \frac{1}{2} (1 + \alpha_X^2 - 4\alpha_M^2); \quad (1)$$

$$\begin{bmatrix} \operatorname{Re}(Z_X) \\ \operatorname{Im}(Z_X) \end{bmatrix} = \begin{bmatrix} \alpha_B & -\alpha_A \\ \alpha_A & \alpha_B \end{bmatrix} \cdot \begin{bmatrix} \operatorname{Re}(Z_S) \\ \operatorname{Im}(Z_S) \end{bmatrix}$$

The model is strongly nonlinear and difficult to treat analytically even in its basic form. In practice, a more refined model taking parasitic effects into account is normally used. An implementation of the method, which defines both impedances as four terminal-pair coaxial standards, has been developed at IEN [3].

3. Numerical analysis

For the present analysis, we considered the simplified model (1) in which the uncertainty contributions are:

- noise in $|U_X|$, $|U_S|$, and $|U_M|$, assumed Gaussian and uncorrelated;
- uncertainty of Z_S , assumed bivariate uniform.

We assigned *a priori* values to Z_S , Z_X and I and $u(|U|)$ to obtain a sample of $3 \times n$ ($n = 40$) simulated readings $\{|U_X|, |U_S|, |U_M|\}_k$ ($k = 1 \dots n$). On the same dataset we performed two different numerical uncertainty analysis.

3.1. Monte Carlo simulation

“Experimental” means and stds from the sample were used as parameters of three Gaussian distributions, from which m Monte Carlo samples were generated (each made of $3 \times n$ voltage values).

To each $3 \times n$ sample, one systematic contribution taken at random from the bivariate uniform pdf of Z_S is associated. Each individual set of three voltages, and the extracted Z_S value (the same for each set), is then processed in the model. This technique is preferred to that suggested in the forthcoming GUM Supplement 1 [4], due to the high nonlinearity of the model. On the n resulting Z_X values the chosen estimator (mean) is computed.

The result is a set of m estimates of Z_X , from which the bivariate pdf is constructed and confidence domains can be computed. In practice, from the Z_X set, we obtained the 95% shortest intervals of confidence for $\text{Re}[Z_X]$ and $\text{Im}[Z_X]$.

3.2. Bootstrap resampling

In this case, the analysis is the same of Par. 3.1., except that, following a prior investigation [5], from the original simulated sample we obtained m bootstrap resamplings (each of $3 \times n$ voltage values).

4. Results

Table 1 shows the results of both methods, for a comparison between a resistor $Z_s = (1000 + j10) \Omega$ with $u(\text{Re}[Z_s]) = u(\text{Im}[Z_s]) = 5 \text{ m}\Omega$, and an inductor $Z_x = (80 + j200\pi) \Omega$, using $I = 2 \text{ mA}$, $u(U) = 10 \mu\text{V}$, $m = 40000$. For brevity only full interval widths $W(\bullet)$ are given (corresponding coverage factors can be computed as $k = W/(2s)$). Fig. 2 shows an example of output bivariate distribution for Z_x .

Table 1. Comparison between MCS and BR, for various combinations of input uncertainties (values given in the text). $s(\bullet)$ is the standard deviation, $W(\bullet)$ is the width of the 95% shortest interval of confidence.

Case	$u(Z_s)$	$u(U)$	$s(\text{Re}[Z_x])$ m Ω	$W(\text{Re}[Z_x])$ m Ω	$s(\text{Im}[Z_x])$ m Ω	$W(\text{Im}[Z_x])$ m Ω
MCS	Yes	Yes	3.7	13.8	3.3	11.6
BR	Yes	Yes	3.6	13.4	3.3	11.6
MCS	No	Yes	2.0	7.8	0.9	3.6
BR	No	Yes	1.9	7.4	1.0	3.9
MCS	Yes	No	3.2	10.5	3.2	10.5
BR	Yes	No	3.2	10.5	3.2	10.5

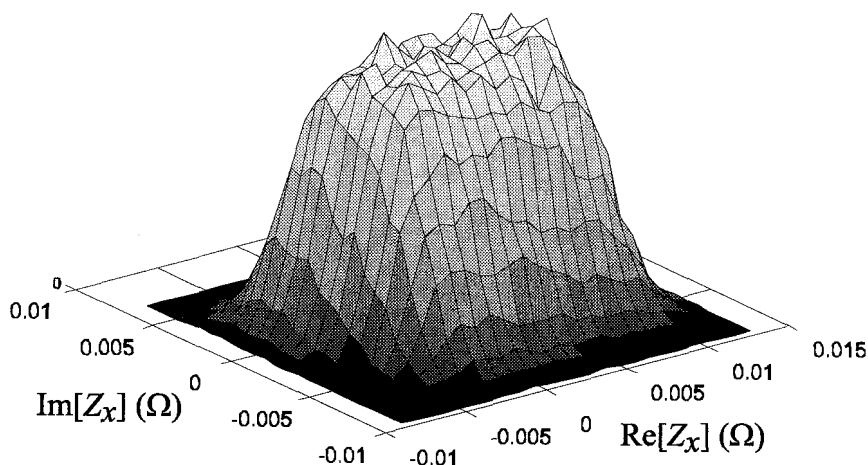


Figure 2. Output bivariate distribution for Z_x (origin in the estimate).

5. Comments

Both methods are able to construct multivariate distributions for vector-valued measurands, with arbitrary models.

The MCS implementation here proposed slightly differs from that suggested in [4]. However, although it is computationally heavier, we believe that it is more suitable for highly nonlinear models.

As concerns the differences between MCS and BR, the latter is more respectful of the experimental information, whereas the former uses statistics depending to some extent on the experimenters' judgement (type of input pdfs, evaluation of input correlations). For example, BR automatically takes into account possible correlations between input data.

BR tends to slightly underestimate intervals of confidence even with sizeable samples ($n=40$ in our case) except in the obvious case in which there is no experimental noise, and the two methods degenerate in one (last two rows of Tab. 1).

In conclusion, we think that BR deserves consideration for those experimental situations, frequent in automated measurements, in which large samples can be obtained. For small sample sizes, MCS should be the preferred choice.

References

1. ISO *Guide to the expression of uncertainty in measurement*, Geneva, 1993-1995.
2. L. A. Marzetta, *IEEE Trans. Instr. Meas.* vol. **IM-21** (4), pp. 353-357 (1972).
3. L. Callegaro, V. D'Elia, *IEEE Trans. Instr. Meas.* **50** (6), pp. 1630-1633 (2001).
4. JCGM/WG1, *Guide to the expression of uncertainty in measurement, Supplement 1: Numerical methods for the propagation of distributions*, draft V5. See also http://www.npl.co.uk/ssfm/download/documents/metrologie_paper.pdf.
5. L. Callegaro, W. Bich, 11th IMEKO TC-4 Symp, Sep. 13-14, 2001, Lisbon, Portugal, pp. 481-484.

BAYESIAN APPROACH TO QUANTUM STATE TOMOGRAPHY*

S. CASTELLETTO, I. P. DEGIOVANNI, M. L. RASTELLO AND I. RUO
BERCHERA

Istituto Elettrotecnico Nazionale "G. Ferraris"
Strada delle Cacce 91, 10135 Torino - Italy

In this paper we focus on the main problem of quantum state tomography: the reconstructed density matrices often are not physical because of experimental noise. We propose a method to avoid this problem using Bayesian statistical theory.

The emerging field of quantum information science exploits quantum mechanics to achieve information processing tasks¹ that are impossible in the classical world. In order to complement and strengthen these technologies, standards and measurement methods must be developed in the realm of quantum metrology. In particular, a central role for benchmarking these quantum information technologies will be played by quantum states tomography (QST)² and quantum process tomography². In this paper we focus on QST, and in particular we propose an estimation method based on Bayesian statistical theory³ to obtain a properly defined reconstructed density matrix (DM).

The more general DM of a quantum system in an N -dimension Hilbert space can be expressed as $\hat{\rho} = (1 + \sum_{i=1}^{N^2-1} s_i \hat{\Omega}_i)/N$, where the $N^2 - 1$ Stokes parameters s_i are real. $\{\hat{\Omega}_i\}$ is a set of $N^2 - 1$ hermitian matrices with null trace ($\text{Tr}\{\hat{\Omega}_i\} = 0$) satisfying $\text{Tr}\{\hat{\Omega}_i \hat{\Omega}_j\} = N\delta_{i,j}$. In order to have a properly defined DM the coefficients s_i satisfy the inequality $\sum_{i=1}^{N^2-1} s_i^2 \leq N - 1$; thus the coefficients s_i can be considered as the coordinates of points in a hyper-sphere (H-S) in $N^2 - 1$ dimensions of radius $\sqrt{N - 1}$. Each point of the H-S corresponds to a possible DM: the points on the surfaces are pure states, and the center of the H-S is the completely mixed state (this is the generalization of the Bloch sphere for a two-dimension Hilbert space¹).

*Work partially funded under EU SofTools_MetroNet Contract N. G6RT-CT-2001-05061

The QST process is based on the ability to reproduce a large number of identical states and perform a series of measurements on complementary aspects of the state within an ensemble in order to obtain the DM of the quantum state from a linear transformation of experimental data². In this case we consider a set of $N^2 - 1$ projective measurement operators $\{\hat{P}_m\}$; thus the probability of observing the measurement outcome m is $p_m = \text{Tr}\{\hat{P}_m \hat{\rho}\}$. As the \hat{P}_m are hermitian with trace one, we can write $\hat{P}_m = \mathbf{1}/N + \sum_{n=1}^{N^2-1} \Gamma_{mn} \hat{\Omega}_n$, leading to the linear relations between the probabilities of measurement outcome and the Stokes coefficients $p_m = 1/N + \sum_{n=1}^{N^2-1} \Gamma_{mn} s_n$. By a proper choice of the measurement operators \hat{P}_m , the matrix Γ_{mn} can be inverted and the Stokes coefficient can be obtained directly from the probabilities of outcomes as $s_n = \sum_{m=1}^{N^2-1} \Gamma_{nm}^{-1} (p_m - 1/N)$. However a significant drawback of this method is that the recovered DM might not correspond to a physical DM because of experimental noise. For example, The DM of any quantum system must be hermitian, positive semi-definite with unit trace; nevertheless the DM reconstructed by QST often fails to be positive semi-definite.

In this paper we study the manipulation of experimental data from a quantum experiment. As quantum theory is a set of rules allowing the computation of probabilities for the outcomes of tests which follow a specified preparations¹, Bayes statistical theory appears to be particularly suitable for this data manipulation. We link the Bayesian notion of prior probability to the information about the quantum state preparation^{1,3}. This information attaches some *a priori* physical constraints allowing proper physically valid DM reconstruction by QST. In particular if we do not have any prior belief the prior probability distribution (Pri-PD) of our system is a uniform distribution in the Bloch H-S $P(\{s_i\}) = V^{-1} \Theta(N - 1 - \sum_{i=1}^{N^2-1} s_i^2)$, where Θ is the step-function and V is the volume of the H-S of radius $\sqrt{N-1}$. The posterior probability distribution³ (Pos-PD) is determined by means of the experimental results according to $P(\{s_i\}|\{E\}) = \frac{L(\{E\}|\{s_i\})P(\{s_i\})}{\sum_{\{s_i\}} L(\{E\}|\{s_i\})P(\{s_i\})}$, where $\{E\}$ is the set of experimental results and $L(\{E\}|\{s_i\})$ is the likelihood function of the experimental results giving the Stokes parameters s_i . The point estimator of s_i is often taken to be the mean of the Pos-PD of s_i ; in this way we always obtain the properly defined DM. This estimator of s_i and the maximum likelihood estimator converge asymptotically to the same value and they are equally efficient. We have to consider that our estimation comes from experimental data as well as from our prior belief. A prior belief corresponding to the physical situation may bias the

estimation. For this reason it is suggested to perform a sequence of measurements where the Pos-PD of a measurement is used as the Pri-PD of the subsequent measurement. We performed some numerical simulation in the case of pure and mixed states of an Hilbert space of dimension 2, using the Adaptive-Genz-Malik algorithm, obtaining an appreciable convergence after only three iteration of the procedure above, even in the presence of biased Pri-PD.

We discuss also the case when the experimentalist needs several copies of a quantum system in a specified state (point Q in the Bloch H-S) to perform a certain experimental task. In practice he is able to prepare the quantum system only approximately in that specified quantum state, but every experimental task may accept a certain tolerance in the quantum state preparation (described by a small H-S centered in Q). We exploit the Bayesian approach to test whether the prepared quantum states is in the intersection between the Bloch H-S and the small H-S Q . We refer to hypothesis Q when we consider the Pri-PD $\pi(\{s_i\})$ (a step-function that is nonzero in the intersection between the two H-Ss), and as the alternative hypothesis A we choose the Pri-PD $P(\{s_i\})$. We define the Bayes factor \mathcal{B} as $\mathcal{B} = \frac{\sum_{\{s_i\}} L(\{E\}|\{s_i\})P(\{s_i\})}{\sum_{\{s_i\}} L(\{E\}|\{s_i\})\pi(\{s_i\})}$; $\mathcal{B} \leq 1$ is in favour of Q against A , while $\mathcal{B} > 1$ is in favour of A against Q . The experimentalist performs the test according to the following procedure: he performs the first measurement and he calculates the parameter \mathcal{B} and the Pos-PD using as $P(\{s_i\})$ the uniform probability distribution in the Bloch H-S. Assuming that the results is $\mathcal{B} \leq 1$, he performs a second measurement and he calculates \mathcal{B} and the new Pri-PD using as $P(\{s_i\})$ the Pos-PD obtained from the first measurement. This procedure is applied iteratively until he obtains $\mathcal{B} > 1$. The Pos-PD of this last measurement allows the experimentalist to decide if the prepared states are in the region of acceptable tolerance or not, for example by estimating the Stokes coefficients s_i and checking if they are in the intersection between the two H-Ss.

We are deeply indebted to W. Bich for helpful suggestions.

References

1. D. Bouwmeester *et al.*, The physics of Quantum Information, (Springer, Berlin, 2000), A. Peres, Quantum Theory: Concepts and Methods, (Kluwer Academic Publishers, Dordrecht, 1995).
2. D.F.V. James *et al.*, *Phys. Rev. A* **64**, 052312 (2001), J.B. Altepeter *et al.*, *Phys. Rev. Lett.* **90**, 193601 (2003), and *ref.s therein*.
3. S.J. Press, Bayesian Statistics: Principles, Models and Applications, (John Wiley and Sons, New York, 1989).

SIMULATION OF CHARGE TRANSFER IN A TUNNEL JUNCTION: APPROACHING THE SUB- e SCALE*

G. E. D'ERRICO

*Istituto di Metrologia "G. Colonnetti", IMGC-CNR, Strada delle Cacce 73, 10135
Torino, Italy, E-mail: g.derrico@imgc.cnr.it*

Based on the orthodox theory of single electronics, a simulation of a tunnel junction is performed, aiming at investigating if quasiparticle events are predictable to transfer fractional charge. The related outcome from the software package MOSES^c (Monte-Carlo Single-Electronics Simulator) is discussed.

1. Introduction

The innovative impact of Single Electron Tunneling (SET) theory and practice is well known in many scientific and technological areas [1]; one field of application is metrology, where many interesting researches are being developed worldwide. [2] The flowering of SET applications involves both the nano- and meso-scales, eventually contributing to a redesign of perspectives of investigations into the constituents of matter too. It is known that findings at the intra-atomic level are jeopardising electrons on a charge of non-elementary particles: two decades ago, quasiparticles were proposed to carry fractional charge and discussed by Laughlin; [3] such quasiparticles have now been successfully investigated in a very recent work, with application to quantum Hall systems.

This paper is concerned with the so-called orthodox theory [5] of single electronics, based on a model where the electric charge is a continuous quantity—constraints on granularity being reduced at a sub- e scale, where electrical shielding phenomena rise at the Debey length. The intended goal is to investigate by simulation if also quasiparticles (instead of electron particles only) are predictable, in the framework of the theory, to transfer fractional charge in a SET device. Emphasis is given on the behaviour of a current-biased, resistively-loaded tunnel junction described in Fig. 1. [5] The research work is developed by use of the software package MOSES^{*} (Monte-Carlo Single-Electronics Simulator). [6] The outcome is discussed, taking into account some side-issues already focused in a related work on computational aspects about counting electrons one by one.

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

^c © 1995 by R.H. Chen ; © 2001 by D.M.R. Kaplan and by Ö. Türel.

2. Simulation: orthodox theory and its application

The orthodox theory is exhaustively expounded by Averin and Likharev: [5] its main result is that single-electron tunneling across a SET device is always a random event, occurring at a rate (its probability per unit time), that depends solely on the reduction of the free electrostatic energy of the device, as required for the event to occur. The process that makes possible the transfer of a single electron despite the crowded junction electrodes (with typically 10^9 free electrons) is a consequence of the Coulomb repulsion of electrons. In order to observe SET effects, temperature constraints are related to the tunnel dimension, approximately: 4.2 K (liquid helium) at 6.0 nm; 77 K (liquid nitrogen) at 1.6 nm; 300 K at 0.8 nm. The theory makes the following major assumptions: the electron energy quantization inside the (metal) conductors is ignored; the time for electron tunneling through the (insulator) barriers is negligible when compared with other time scales; coherent quantum processes responsible of possible simultaneous (co-tunneling) events are ignored. Focusing on the subject of the present research, a dummy assumption that can be made explicit regards the charge flow; if Q_t is the charge transferred across the junction during a given time interval Δt , then (Fig. 1) $Q_t \equiv q - Q$, where q is the time integral of $I_t \equiv I_0 - I_s$:

$$q \equiv \int_{\Delta t} I_t dt. \quad (1)$$

According to Eq. (1), the charge is theoretically allowed to take also any fractional value of e , making events of quasi-particle transfer likely to occur, being in principle constrained by the time interval Δt .

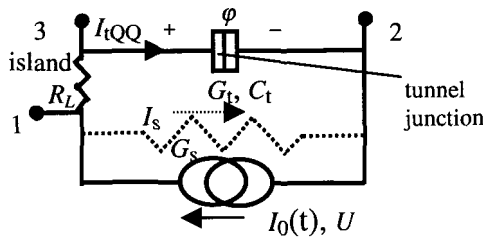


Figure 1. Current-biased, resistively (R_L) loaded tunnel junction: I_t , Q , C_t , and G_t , tunnel current, charge, capacitance and conductance; U voltage input between nodes 1 and 2 (reference); ϕ voltage between island 3 and node 2; I_s and G_s , shunt current and conductance (ignored in simulation).

Table 1. Simulation response: time-average jumps, island charge Q and voltage ϕ . Three runs data.

U	Jumps			Island charge			ϕ
0.4525	1.03	0.996	1.04	0.431	0.471	0.455	0.436
0.4475	0.732	0.679	0.707	0.393	0.415	0.444	0.428
0.4425	0.503	0.516	0.521	0.379	0.436	0.384	0.420

Fractional charge tunneling in a SET device (Fig. 1) has been simulated, using MOSES*. [7] The simulation parameters are normalized in terms of capacitance C_i and conductance G_i (once the simulation is over, the results can be converted to real units, by introducing the actual values for C_i and G_i). Used units are: charge: e ; energy: e^2/C_i ; voltage: e/C_i ; time: C_i/G_i ; temperature: $e^2/(kC_i)$ (k, Boltzmann constant) e.g., if $C_i=10^{-16}$ farad, the unit temperature is 18 K. In our full Monte-Carlo simulation all input data are set to 1, but temperature 0.001, and single run iterations are performed with: time step: 10^{-1} , number of periods per run: 10^3 , max total time per run: 10^6 , time period: 10^3 . After exploratory trials, the dc voltage is tuned below $U=0.4525$, corresponding to single whole- e jumps across the tunnel. Tab. 1 shows some results, including time-averaged sub- e jumps.

3. Final remarks

Fractional jump values, relative to the net jump of one particle e , can be interpreted as the tunneling of corresponding sub- e particles. The simulation result is a criticism to electron individuality, opening sub- e perspectives too.

References

1. K.K. Likharev, *Proc. IEEE*, **87**, 606 (1999).
2. K. Flensberg, A.A. Odintsov, F. Liefink, P. Teunissen, *LANL e-Print Archive*, [arXiv:cond-mat/9908219](http://arxiv.org/abs/cond-mat/9908219) (16 August, 1999).
3. R.B. Laughlin, *Phys. Rev. Lett.*, **50/18**, 1395 (1983).
4. S. Roddaro, V. Pellegrini, F. Beltram, G. Biasiol, L. Sorba, R. Raimondi, G. Vignale, *Phys. Rev. Lett.*, **90/4**, 046805–1 (2003).
5. D.V. Averin, K.K. Likharev, in: *Mesoscopic Phenomena in Solids* (B.L. Altshuler, P.A. Les, R.A. Webb, eds.), chap. 6, Elsevier Science Publ., Amsterdam, 1991.
6. <http://hana.physics.sunysb.edu/set/software/index.html>, *Moses User Guide (1.1 for DOS)*, State University of New York at Stony Brook (1995-2001).

VALIDATION OF CALIBRATION METHODS - A PRACTICAL APPROACH*

EDUARDA FILIPE

Instituto Português da Qualidade (IPQ)

Rua António Gião, 2

Caparica – Portugal

E-mail: efilipe@mail.ipq.pt

Abstract

To comply with point 5.4.5 of ISO/IEC 17025 standard means that the laboratories “shall validate non-standard methods, laboratory designed/developed methods, standard methods used outside their intended scope, and amplifications and modifications of standard methods to confirm that the methods are fit for the intended use” [1]. This requisite of the standard is new and the laboratories are evaluating the approaches for the validation process in terms of the quality system. A procedure for validation of calibration methods is proposed and an example of validation of results of measurements is described.

1. Introduction

To fulfil the requisites of the ISO/IEC 17025 standard [1] (General Requirements for the Competence of Testing and Calibration Laboratories) means that the laboratories define their quality objectives, the performance of the tests, reference materials certifications and calibrations are in agreement with the methods described in technical procedures, with the clients requirements and with the use of the good professional practices; the validation of methods and results is performed; the fulfilment of the objectives of the activity plans is revised by the management; the performance of the processes are evaluated, as the audits reports, the satisfaction inquiries, the information from the complaints, preventive and corrective actions are carried on and the outputs from the working groups for improvement are considered.

A brief overview of the validation concept will be considered in this text and the fulfilment of point 5.4.5 of the ISO/IEC 17025 standard. A procedure for validation of calibration methods is proposed and an example of validation of results of measurements is described using the statistical technique

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

‘Experimental Design’, active tool that enables inferences about the results, namely from reproducibility measurements and methods comparisons.

1.1 General principles and concepts

The ISO standard 9000:2000 [2] defines "validation" as "confirmation through the provision of objective evidence that the requirements for a specific intended use or application have been fulfilled...^(c)".

The ISO IEC 17025 states at point 5.4.5 that ‘the laboratory shall validate non-standard methods, laboratory designed/developed methods, standard methods used outside their intended scope, and amplifications and modifications of standard methods to confirm that the methods are fit for the intended use’.

As a consequence the laboratories should describe the way the validation of calibration methods is done and this description should be a part of its quality documents. In the validation process, the method is evaluated in terms of its representativeness, repeatability and reproducibility^(d) and considered its ability to satisfy the defined requirements, its technical capability and the expected or required uncertainty of the results.

2. Procedures for method validation

The Laboratory Quality Manual will state the general principles to observe in the validation of the calibration methods. Each calibration procedure will include also a validation chapter that describes the way the validation is performed, the calibration requirements and how to fulfil these requirements, and the statistical tools to be used to validate the results. To accomplish these goals, studies or developments will be performed, taking in consideration if the methods are internal, new or used by other laboratories, or reference methods.

At EUROLAB Report [3] two major approaches are considered, for the validation process, the scientific approach and/or the comparative one.

In the first approach, the scientific approach, the method representativeness is evaluated by the published scientific and technical literature, by the research and developments performed by the laboratory, by simulation and modelling, by

^c In the ISO 9000:2000 are this two notes: Note 1: The term "validated" is used to designate the corresponding status and Note 2: The used conditions for validation can be real or simulated.

^d (VIM, 3.6 and 3.7 [4]) Repeatability (of results of measurement) - Closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement. Reproducibility (of results of measurements) Closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement.

the method robustness and by the studies for the optimisation of the laboratory performance. In the other approach, the comparative one, the method is assessed by comparing its results to those obtained by other validated method, which has been developed for the same purpose. The comparisons may be performed with standard methods, reference methods, international standards or certified reference materials and/or interlaboratory comparisons. In the method comparisons the results must fall within the comparison uncertainty interval.

Finally, the validation of the method will be obtained by the combined use of the above-described procedures.

3. Validation of results

3.1 General Model

Calibration laboratories are also required to validate the results of measurements and the associated uncertainties. The laboratory designs the experiment and an adequate number of measurements are performed. For the results analysis it is proposed the use of one of the statistical tools of the “Experimental design”.

These tools are used generally for the improvement and optimisation of a process. It is an active method that allows through the changes in the inputs and observing the corresponding changes in the outputs, to make the inference by rejecting the *null hypothesis* (H_0) of which outputs are statistically different for a significance level α [5], also known as the “producer’s” risk^(e).

Inversely it can be used to test the homogeneity of a sample^(f), for the same significance level, to identify which results can be considered as outliers and the way to deal with them in a SPC^(g) concept and what is the uncertainty associated to the sample. The procedure to be described is the well-known analysis of variance (*ANOVA*).

3.2 Example of validation of calibration results of a comparison of two thermometric tin fixed point cells.

Short description of the laboratorial work

For the comparison of the two tin sealed cells, we have used two furnaces in order to realize the tin freezing point ($t=231,928\text{ }^{\circ}\text{C}$) simultaneously with each cell in its own furnace. Were performed 9 different freezing point plateaus

^e The probability of making the so-called *type I* error, to reject the null hypothesis and it is true.

^f In this case sample is the set of measurements.

^g SPC – Statistical Process Control

for each cell in consecutive days, and utilised 3 standard platinum resistance thermometers (SPRTs) A, B and C at each plateau, that allowed to obtain, each day, three sets of measurement differences. It was chosen the “randomised block design without interaction” to evaluate if there were effects between the measurements performed at each day and/or between the measurements performed by each thermometer.

If no effects are detected the total average of the differences of the two cells will be a good estimator of the difference of the cells.

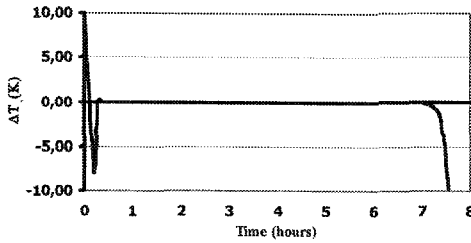


Figure 1: Freezing point plateau of a tin fixed-point

The model

If we have different effects that we wish to compare, each group of observations for each effect is a random variable. It is possible to describe the observations by the model:

mn observations

$$Y = [y_{ij}] \quad (i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n)$$

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

where y_{ij} is the $(ij)^{\text{th}}$ observation, μ the overall mean, τ_i the i^{th} factor-line effect (also known as treatment), β_j the j^{th} factor-column effect (also known as block) are considered as deviations from the overall mean and ε_{ij} the random error component, normally and independently distributed with mean zero and variance σ^2 . The variance is assumed to be constant.

To test the null hypotheses:

$$H_0^{\text{Line}}: \tau_1 = \tau_2 = \dots = \tau_m = 0 \quad \text{and} \quad H_0^{\text{Column}}: \beta_1 = \beta_2 = \dots = \beta_n = 0$$

We use the statistics:

$$F_0^{\text{Line}} = \frac{MS_{\text{Line}}}{MS_{\text{Error}}} \sim F_{\alpha, m-1, (m-1)(n-1)}$$

$$F_0^{\text{Column}} = \frac{MS_{\text{Column}}}{MS_{\text{Error}}} \sim F_{\alpha, n-1, (m-1)(n-1)}$$

where: F – sampling distribution ^(h); MS – mean square between factors (line or column) is the unbiased estimate of the variance σ^2 if H_0 is true, sureestimate of σ^2 if it is false and MS_{Error} – mean square error is always an unbiased estimate of the variance σ^2 .

If $F_{0^{Line}} > F_{\alpha, m-1, (m-1)(n-1)}$ and $F_{0^{Column}} > F_{\alpha, n-1, (m-1)(n-1)}$, the null hypotheses are rejected for a significance level α and the results are statistically different. If the above conditions are not true, the null hypotheses are not rejected for a significance level α , the results are “not different” and the samples belong to the same statistical population. Table 1 displays the 27 obtained temperature

TABLE 1
Tin Freezing point - Comparison of two cells

		Temperature differences ΔT (mK)								
		Days	1	2	3	4	5	6	7	8
SPRT	A	0,12	0,07	0,11	0,07	0,05	0,02	0,00	0,05	0,07
	B	0,07	0,06	0,09	0,09	0,01	0,00	0,05	0,07	0,05
	C	0,06	0,08	0,06	0,15	0,06	0,00	0,06	0,14	0,12

differences between the two cells (see also Fig. 2). From the ANOVA Table 1 we have obtained two F_0 values that will be compared with the F distribution for $\alpha = 5\%$ and 2 and 16 degrees of freedom, $F_{0,05, 2, 16} = 3,6337$ for the SPRT effect and F distribution for $\alpha = 5\%$ and 8 and 16 degrees of freedom $F_{0,05, 7, 16} = 2,5911$ for the measurements/days effect.

The null hypothesis is rejected for the measurements / days effect but not rejected for the measurements performed by the three thermometers.

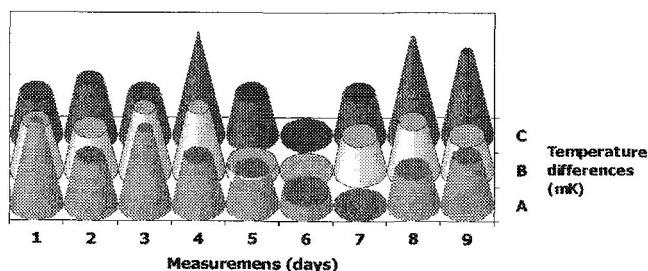


Figure 2: Schematic representation of the observed temperatures differences

^h F distribution – Sampling distribution. If χ_u^2 and χ_v^2 are two independent chi-square random variables with u and v degrees of freedom, then its ratio $F_{u,v}$ is distributed as F with u numerator and v denominator degrees of freedom.

3.3 Criteria for dealing with outlying measurements

It is proposed to use a sample statistic to be compared with a critical value in order to determine the measurements to reject.

ANOVA TABLE I				
	Sum of squares	Degrees of freedom	Mean Square	F ₀
Factor-line	0,0034	2	0,0017	1,8787
Factor-column	0,0235	8	0,0029	3,2539
Error	0,0144	16	0,0009	
Total	0,0413	23		

From ANOVA calculations we have detected the existence of a day or days where the values may be considered as outliers.

The method to detect these measurements was the “Test on the mean of a normal distribution with unknown variance” for a significance value of 5%. The normality assumption is needed to formally develop the test but “moderate departures for normality will not seriously affect the results” [5, pp.79].

The test statistic used is $t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ and the null hypothesis will be rejected if $|t_0| > t_{\alpha/2, n-1}$ where $t_{\alpha/2, n-1}$ is the upper $\alpha/2$ (bilateral test) of the t -distribution with $(n-1)$ degrees of freedom.

The average of the 27 measurements will be considered as the mean μ_0 and equal to 0,066 mK. The 9 daily averages will be used to calculate the $t_{0, day}$.

Table 2 drafts these values. Comparing them with $t_{0,025, 8} = 2,7515$ we conclude that the measurements from the day 6, are to be rejected.

TABLE 2					
Day	1	2	3	4	5
$t_{0, day}$	0,9645	0,2257	1,1492	2,0727	1,4365
Day	6	7	8	9	
$t_{0, day}$	3,2835	1,6212	1,1492	0,7798	

Now we recalculate the ANOVA Table omitting the values from day 6. The obtained F_0 values will be compared again with the F distribution for $\alpha = 5\%$ and 2 and 14 degrees of freedom, $F_{0,05, 2, 14} = 3,7389$ for the SPRT effect and F distribution for $\alpha = 5\%$ and 7 and 14 degrees of freedom $F_{0,05, 7, 14} = 2,7642$ for the measurements/days effect.

The null hypotheses is not rejected, so there is no effect between the measurements performed with the different thermometers or between the measurements/ days, for a significance level $\alpha = 5\%$.

ANOVA TABLE 2

	Sum of squares	Degrees of freedom	Mean Square	F ₀
Factor-line	0,0040	2	0,0020	2,0746
Factor-column	0,0116	7	0,0017	1,7153
Error	0,0135	14	0,0010	
Total	0,0291	23		

Alternatively, it may be stated that there is no significant difference between the 24 sets of measurements, they belong to the same statistical population and the average ($\Delta T = 0,073$ mK) and the variance of the mean ($\sigma^2_{\Delta T} = 5,28 \times 10^{-5}$ mK²) with 23 degrees of freedom are good estimates, respectively, of the mean and the variance of the differences of the two fixed point cells.

We are able now “to provide objective evidence, that the requirements for a specific intended use or application have been fulfilled...” what is to say that we have performed the validation of the comparison measurement results.

As a consequence, the obtained standard deviation of the mean ($\sigma_{\Delta T} = 0,0073$ mK) will be appropriate for the evaluation of type A component of uncertainty of the comparison. The type B components, from prior information will be added in order to complete the uncertainty budget.

4. Conclusion Remarks

It was outlined an approach to answer to point 5.4.5 of the new ISO/IEC 17025 standard and described what should be included in the Quality documents for the validation of the methods and results.

An application of the randomised block design has been used to illustrate a practical realization of validation of a comparison of two thermometric fixed points results in repeatability/reproducibility situation that can be easily applied to the calibration results.

References

1. ISO, IEC - General Requirements for the Competence of Testing and Calibration Laboratories. International Organization for Standardization,

International Electrotechnical Commission, ISO IEC 17025:2000, Genève, 2000.

2. ISO - Quality management systems: Fundamentals and vocabulary. International Organization for Standardization, ISO 9000:2000, Genève, 2000.
3. EUROLAB – “Validation of Test Methods. General Principles and Concepts”. European Federation of National Associations of Measurement, Testing and Analytical Laboratories. Internal Report EL1545/96, Brussels, 1996.
4. BIPM et al - International vocabulary of basic and general terms in metrology (VIM). International Organization for Standardization, Genève, 1993. ISBN 92-67-01075-1.
5. Montgomery, D. – Introduction to Statistical Quality Control. Singapore: John Wiley & Sons, 2nd ed., 1990. ISBN 0-471-51988-X.
6. Murteira, B. – Probabilidades e Estatística. Volume II. Amadora: Mc-Graw Hill de Portugal, 2nd ed., 1990. ISBN 972-9241-07-4.

COMPARISON OF LS TECHNIQUES FOR THE LINEAR APPROXIMATION OF DATA AFFECTED BY HETEROSCHEDASTIC ERRORS IN BOTH VARIABLES WITH UNCERTAINTY ESTIMATION*

D. ICHIM¹, P. CIARLINI², F. PAVESE¹, A. PREMOLI³, M.L. RASTELLO⁴

¹ CNR, Istituto di Metrologia “G.Colonnetti”, strada delle Cacce 73, 10135 Torino, Italy
E-mail: d.ichim@imgc.cnr.it

² CNR, Istituto per le Applicazioni del Calcolo “M.Picone”, viale del Policlinico 137,
00161 Roma, Italy. E-mail: ciarlini@iac.rm.cnr.it

³ Politecnico di Milano, p.le Leonardo da Vinci 32, 20133 Milano, Italy.

⁴ Istituto Elettrotecnico Nazionale “G.Ferraris”, strada delle Cacce 91, 10139 Torino,
Italy. E-mail: rastellorastello@ien.it

Various least squares methods have been compared in respect to the straight line fitting to data sets with errors on both variables, to check the benefit in using the most appropriate method for dealing with heteroschedastic data, the element-wise total least squares (EW TLS). It is found that the EW TLS always gives the correct estimate: weighted least squares can sometimes be also a good approximation, but this cannot be guessed *a priori*.

1. Introduction

When the experimental data are affected by errors in both variables, the metrologist faces the problem that the ordinary least squares approximation (OLS) does not provide an optimal answer in estimating the regression parameters and their uncertainty. The same situation arises when the data are, additionally, heteroschedastic and the weighted least squares approximation (WLS) is used, since the simple weighting operation does not adjust for errors in the independent variable.

In an OLS problem the minimization of sum of squared distances along one coordinate axis is performed, while TLS aims to minimize sum of squared perpendicular distances to the best line and the TLS correction is always smaller than the one obtained with OLS. In the present work a method, recently developed by two of the authors (AP and MLR) called element-wise TLS (EW TLS), has been considered. It uses an iterative procedure to approximate a TLS solution for the heteroschedastic case [1]. Since it does not presently allow a

* Work partially funded under EU SofTools_MetroNet Contract N. G6RT-CT-2001-05061

direct estimation of the model parameter uncertainty, its results have been compared with the OLS and the WLS ones.

Two series of data has been considered in this testing. The first is a set of real data under study in thermal metrology, where the range of the uncertainty values is quite broad for the independent variable. The second series is a simulated set of data constructed in order to emphasise the benefit of the use of the correct method, the TLS.

The uncertainty of the model parameters has been evaluated by means of a bootstrap technique [2], and the overall uncertainty with a Monte Carlo technique [3] taking into account the uncertainty on both variables of each experimental point, both in the cases of OLS and WLS.

2. Comparison of LS Evaluations with Real Thermal Data Series

Table 1 reports the set of data. The range of uncertainties is much broader for one variable than for the other. In addition, the data arise from different series of independent measurements: the position of the zero is arbitrary and the relative position has been set visually. Therefore, they have been considered either as an overall set or as subdivided in the original series, in order to additionally detect with the fixed-effect model technique [4] a *translation* effect or, equivalently, possible biases between series when using OLS or WLS.

The main characteristics of the data are:

- different $U(\text{Lab})$ and $U(x)$ from point to point
- unknown relative position Δy between Lab measurements
- one outlying point at visual examination (in grey in Table 1)
- linear ΔT vs x model: $y = i + s x$

Table 1. Thermal data test series.

#	#		y		x		#	#		y		x	
			U(y)		U(x)					U(y)		U(x)	
1	1	Lab1	-75	0,045	27,5	0,2	2	3	Lab2	-60	0,030	27,5	0,2
	2	Lab1	-60	0,045	27,5	0,2		6	Lab2	-15	0,030	33,4	0,2
	4	Lab1	-30	0,045	29,1	0,6		14	Lab2	350	0,030	101,0	0,3
	9	Lab1	20	0,045	45,8	0,6	3	5	Lab3	-30	0,020	29,1	0,6
	10	Lab1	20	0,045	46,2	3,2		7	Lab3	18	0,020	45,8	0,6
	12	Lab1	295	0,045	91,6	0,8		8	Lab3	60	0,020	45,8	0,6
	15	Lab1	370	0,045	103,8	2,0	4	11	Lab4	20	0,107	46,2	3,2
	16	Lab1	460	0,045	123,0	6,0		13	Lab4	300	0,070	91,6	0,8
	17	Lab1	395	0,045	124,0	6,0		19	Lab4	440	0,070	124,0	6,0
	18	Lab1	470	0,045	124,0	6,0		21	Lab4	580	0,070	154,9	1,6
	20	Lab1	635	0,045	154,9	1,6							

The results of the estimation with different LS methods are reported in Table 2. The data of Lab1, which represent about half the set, obviously appear homogeneous, if the outlier is omitted; the omission halves the standard deviation of the fit residuals and also significantly reduces the bootstrap confidence interval. The use of Mathematica® software package gives a slope value, which represents the most important indicator for these data, $s = 5,2$.

Table 2. Results on the thermal data test series.

	OLS	WLS	EW	TLS	OLS	OLS	WLS
	<i>Intercept i</i>				Bootstrap (95%CI)		with fixed effect
					All data		
All Labs	-202	-204	-205		-(184-220)		
s.d.	23	23	24		22		
Lab1	-213		-214		-(186-234)		
s.d.	25		25				
					No outlier		
All Labs	-210	-208	-208		-(189-223)		
s.d.	17	17	17		17		
Lab1	-217		-217		-(198-240)		
s.d.	14		14				
	<i>Slope s</i>				Bootstrap (95%CI)		with fixed effect
					All data		
All Labs	5,26	5,31	5,34		5,01-5,52		
Lab1	5,39		5,43		4,98-5,68		
					No outlier		
All Labs	5,34	5,42	5,42		5,13-5,58		
Lab1	5,53		5,53		5,36-5,76		

The intercept value estimates show no significant difference (range of 15) from OLS to EW TLS estimation, compared with the bootstrap 95%CI interval (34-48). However there is a decrease of about $\approx 5\%$ when the outlier is discarded, comparable to the bootstrap 67%semi-CI (0,08-0,12).

The slope estimates show a range 5,26–5,53, i.e. a variation of 0,27, but a definite increase of the slope of about 0,1-0,2 when the outlier is omitted, comparable with the bootstrap 67% semi-CI (0,10-0,17). By using the WLS almost the same results of the EW TLS are obtained. For WLS, it may be significant to apply it also to the reversed variables: in this case, one gets a set of new slope values lying in the same range.

The application of the fixed effect method provides the optimal positions of each series with respect to one taken (arbitrarily) as the reference, here Lab1. Only the translation for series Lab4 when the outlier is omitted, $\Delta i = -20$, is statistically significant, being higher than the standard deviation of the fit. The slope and intercept values lie within the previous ranges.

An overall estimate of the uncertainty has been obtained with a Monte Carlo simulation approach: 1000 sets of data are randomly sorted from the bidimensional *rectangular* pdf defined by the 95%CI of either variable; the

parameters obtained by fitting the simulated datasets are used to infer their uncertainty. The best results (Fig.1) give an uncertainty $\varepsilon_{WLS, \min}(y) = 32,6$ occurring at $x = 55 \pm 3$, corresponding to about half the fit best standard deviation.

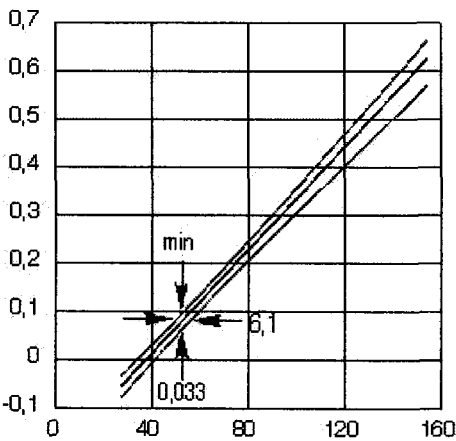


Figure 1. 95%CI boundaries of the Monte Carlo simulation using WLS and no outlier.

3. Comparison of LS Evaluations with Simulated Data

Table 3 reports the set of simulated data (single series), also based on real metrological data. In Table 4 are the results with different methods.

Table 3. Simulated data.

#	y		x		#	y		x	
		<i>U(y)</i>		<i>U(x)</i>			<i>U(y)</i>		<i>U(x)</i>
1	1,0	0,28	0,992	0,28	9	9,0	0,18	8,9973	0,18
2	2,0	0,30	2,0076	0,30	10	10,0	0,50	10	0,50
3	3,0	0,18	3,0027	0,18	11	11,0	0,25	11,1122	0,25
4	4,0	0,30	3,986	0,30	12	12,0	0,28	11,9556	0,28
5	5,0	0,25	4,9595	0,25	13	13,0	0,17	13,0455	0,17
6	6,0	0,30	6,126	0,30	14	14,0	0,77	13,9664	0,77
7	7,0	0,12	7,0091	0,12	15	15,0	0,24	15,045	0,24
8	8,0	0,20	7,988	0,20	16	16,0	0,25	15,9744	0,25

Table 4. Results on simulated data (fit s.d. = 0,050).

OLS	WLS	EW TLS
	<i>Intercept i</i>	
0,00741	0,01550	0,01159
	<i>Slope s</i>	
1,00036	0,99987	0,99753

In this case, the WLS is only marginally different from OLS. On the contrary, the use of EW TLS, the mathematically accurate method, gives significantly different results, as compared with the fit uncertainty: the intercept is about 2σ lower than the OLS and WLS estimates. Monte Carlo simulation estimates are significantly better for WLS than for OLS: $\varepsilon(y)_{OLS} = 0,42$ at $x = 7,46 \pm 0,21$ (Fig.2). The value $\varepsilon_{WLS,min}(y) = 0,33$, at $x = 7,98 \pm 0,16$, corresponds to the fit standard deviation.

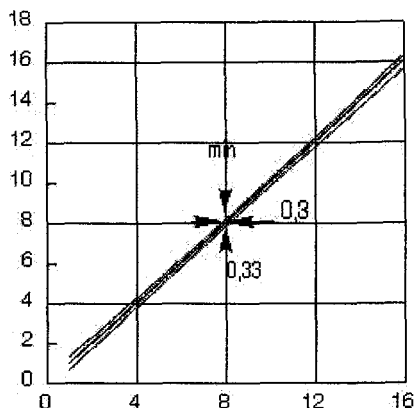


Figure 2. 95%CI boundaries of the Monte Carlo simulation using WLS.

4. Conclusions

The comparison of different LS methods applied to the case of data with heteroschedastic errors on both variables have shown that the WLS may be sometimes sufficient, depending on the type of weighting scheme with respect to the EW TLS method, which is a more computer-intensive method, but more mathematically accurate. Consequently, WLS could be sometimes sufficient, after a rational comparison and discussion made by performing several least-squares estimates. A Monte Carlo simulation using the uncertainty CIs of the data on both variables, as pdf can be a powerful tool in providing an estimate of the overall uncertainty of the fitting model.

References

1. M.L. Rastello and A. Premoli: Homotopic solution of EW-TLS problems, *This book*.
2. P.Ciarlini, G.Regoliosi and F.Pavese: "Non-parametric bootstrap with applications to metrological data", in: "Advanced Mathematical Tools in Metrology, vol.1", Series on Advances in Mathematics for Applied Sciences vol.16, World Scientific, Singapore, 1994, 219-230.

3. D.Ichim, F.Pavese, J.Casas-Cubillos and C.Balle, *Temperature, Its Measurements and Control in Science and Industry* (D.Ripple Editor), AIP, New York, 2002, 429-33.
4. F.Pavese and P.Ciarlini: "Accurate modelling of translational bias and its application to reduction of thermophysical data series", *Metrologia* 27 (1990), 145-152.

NOISE CORRECTION FOR SURFACE MEASUREMENTS

H. HAITJEMA AND M.A.A. MOREL

*Precision Engineering section,
 Eindhoven University of Technology, PO Box 513,
 5600 MB Eindhoven, The Netherlands
 E-mail: : m.a.a.morel@tue.nl*

In roughness measurements often noise plays a role. Noise may give an offset in measurement parameter as noise makes the parameter deviate away from zero. In this paper we propose a method to correct for noise bias for the surface parameter Sq . By considering the decrease in Sq once an average over multiple measurements is made, an unbiased value for Sq is estimated by extrapolating the value to an infinite amount of measurements. It is shown that using this method for two measurements only, the true measurand is approached better than with averaging tens of measurements. This principle is extended to obtain a complete 'noise-corrected' surface by considering the power spectrum and the change of each Fourier component with averaging. Combining the two methods and considering the statistical significance of each Fourier component enables a further reduction. Examples and simulations are shown for the calibration of roughness drive axis and surface measurements.

1. Method

1.1. Parameter extrapolation

The parameter Sq in surface metrology is a measure of power as it is the rms value of the surface. It is possible to use the mean square, Sq^2 to extrapolate the Sq to a power level with reduced noise bias. As the noise is independent of the measurement, the power of a measurement consist of the power of the object and the power of noise [1]:

$$Sq_{\text{measurement}}^2 = Sq_{\text{object}}^2 + Sq_{\text{noise}}^2 \quad (1)$$

If multiple measurements are taken, noise is reduced by the number of measurements n , but the power of the object remains:

$$Sq_{\text{mean}(n)}^2 = Sq_{\text{object}}^2 + Sq_{\text{noise}}^2 / n \quad (2)$$

Rearranging 2 and 3 gives the power of the object:

$$Sq_{\text{object}}^2 = \frac{n \cdot Sq_{\text{mean}}^2 - \langle Sq_{\text{measurement}}^2 \rangle}{n - 1} \quad (3)$$

In figure 1, the results are given of the extrapolation for a simulated block surface with an Sq of 1.381 μm . Measurement surface is 0.01 mm^2 , period 10 μm and amplitude 1 μm . A normal distributed noise with a standard deviation of

1 μm is added. As the standard uncertainty in the extrapolated sets does not decrease, e.g. averaging two profiles gives $Sq = 1.547 \mu\text{m} \pm 12 \text{ nm}$ and the reduced pairs give $Sq = 1.376 \mu\text{m} \pm 10 \text{ nm}$, it is possible the calculated parameter is lower as the 'clean' surface, but within the uncertainty limits. The bias effect is observed in surface plates [2] and noted for rms values of wavefronts by Davies [3]. Davies uses a combinatorial method to remove the bias.

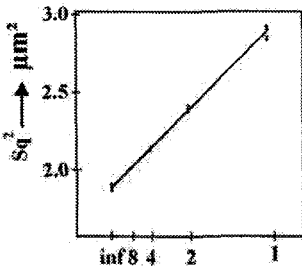
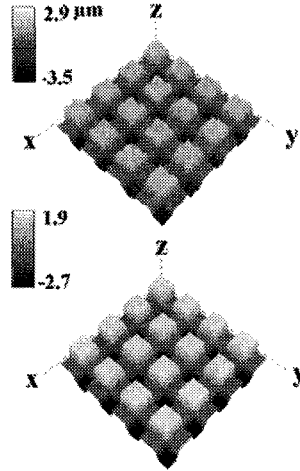


Figure 1.

Extrapolated sets of two measurements compared to the mean of two, four and eight measurements and the single results. The mean of eight files surface is in the upper right ($Sq = 1.419 \mu\text{m}$), the random phase exclusion surface is lower right ($Sq = 1.374 \mu\text{m}$).



1.2. Fourier reduction

The next step is to obtain a complete reduced surface. First a Fourier transformation is performed. The power, the squared amplitude, of a single frequency also consists of an object and a noise part. Following formulas 1 to 3, the amplitude spectrum of a noise reduced surface can be calculated. Inverse transformation is done with the phase of the mean of measurements. If a noisy frequency is encountered it is possible the power is reduced to a negative number. In that case the power of the frequency is set to zero.

1.3. Random phase exclusion

Even further reduction is possible. The Fourier terms are sorted by statistical importance. The more random a frequency, the larger the standard deviation in the mean of phases of that frequency will be and the less significant it is. Starting with the most significant frequency, Sq is calculated and compared to the extrapolated Sq_{object} calculated with equation 3. If the limit is not met the next frequency is added to the spectrum until $Sq_{surface}$ meets Sq_{object} .

2. Calibration

A drive axis of a roughness tester is calibrated with a randomizing method [4]. A reference plain is measured at 6 different positions. The reference and its phase is thus randomized while the axis remains constant. Using the noise correction and random phase exclusion methods, a plot of the axis is derived.

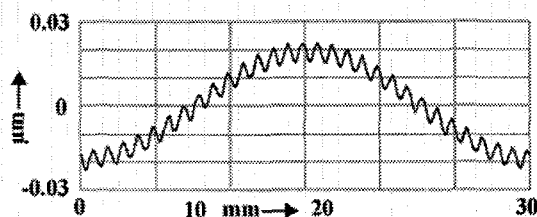


Figure 2. Calibration of drive axis of roughness tester with random phase exclusion.

Acknowledgement

This research is supported by Mitutoyo Nederland B.V.

References

1. C. Chatfield, *Statistics for Technology*, Chapman & Hall 1983, ISBN 0-412-25340-2
2. H. Haitjema, *Uncertainty propagation in Surface Plate measurements*, in ISMQC 4th International symposium on dimensional metrology; Editors: H. Tikka, Tampere, Finland, 304-320, (1992)
3. A. Davies and M.S. Levinson, *Estimating the root mean square of a wavefront and its uncertainty*, *Applied Optics* 40 (2001) pp 6203 - 6209.
4. C.J.Evans, R.J. Hocken and W.T. Estler, *'Self-calibration: reversal, redundancy, error separation and absolute testing'*, *Annals of the CIRP*, Vol 45/2, (1996) pp 617-634.

EVALUATION OF UNCERTAINTY OF STANDARD PLATINUM RESISTANCE THERMOMETER AT NATIONAL LABORATORY LEVEL*

M. J. KORCZYNSKI AND A. HETMAN

*Technical University of Lodz,
Institute of Theoretical Electrotechnics, Metrology and Material
18/22 Stefanowskiego Street, 90-924 Lodz, Poland
E-mail: jerzykor@p.lodz.p*

A. SZMYRKA-GRZEBYK

*Institute of Low Temperature and Structure Research, Polish Academy of Science
2 Okólna Street, 50-422 Wrocław, Poland
E-mail: szmyrka@int.pan.wroc.pl*

P. FOTOWICZ

*Central Office of Measures (GUM)
2 Elektoralna Street, 00-950 Warszawa, Poland
E-mail: gum@gum.gov.pl*

The main purpose of the paper is to present uncertainty evaluation of Standard Platinum Resistance Thermometers (SPRT). The three methods of calculation of expanded uncertainty are presented and the results are compared.

1. The Challenge

The national standard thermometer used for the range 13.8033 K to 273.16 K is calibrated according to International Temperature Scale, ITS-90, at the triple point of: Hydrogen, Oxygen, Argon, Neon, Mercury and Water.

Those temperatures are produced using an appropriate cryostat. Instrumentation associated with the measurement laboratory set up consisted of fixed-point cells, platinum resistance sensors, alternating current resistance bridges, and standard resistors with thermostats and the thermostat for water cell.

The reproduced temperature to be measured is defined by the value of the resistance of SPRT at each fixed point (which corresponds to each of the 5

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

points: Hydrogen, Oxygen, Argon, Neon and Mercury), compared to the resistance of the SPRT at reference point (triple point of Water).

The input quantities considered as a type A, are characterized by probability density functions (pdfs), as t-Student's and Gaussian distributions and type B as uniform pdfs. Three methods for calculating coverage factor [1] at a level of confidence of 0.95 are provided. Two methods are based on calculation of combined standard uncertainty. The first one uses an arbitrary coverage factor while the second uses an approximation of a coverage factor. The third method is based on calculating the convolution of input quantities. The results of the calculations of these three methods are compared for all 5 equilibrium points. In a previous paper [3] some preliminary information is given about applying the procedures described in here to working types of PRTs. By contrast, the present paper focuses upon "national" SPRTs by considering how components contribute to the overall uncertainty, and calculation of uncertainties for these units.

The influence of all external quantities was considered as thoroughly as possible. It can be noticed that although most of these were non-correlated, some of them were. The effects of both the correlated and uncorrelated factors were taken into account in the measurement equation.

2. Equation of measurement

Measurement equation (1) presents the mathematical model of measurand as a quotient of the resistance of the thermometer placed in the medium at any fixed point, e.g. hydrogen triple point, and the resistance of the thermometer placed in the substance at the stipulated reference point.

$$W = \frac{R_{Tc}}{R_{0c}} = \frac{R_T + \delta R_{T1} + \delta R_{T2} + \delta R_{T3} + \delta R_{T4} + \delta R_{T5}}{R_0 + \delta R_{01} + \delta R_{02} + \delta R_{03} + \delta R_{04} + \delta R_{05} + \delta R_{06}} \quad (1)$$

R_{Tc} – resistance at any fixed point with corrections

R_{0c} – resistance at the reference point with corrections

R_T – resistance at any fixed point without corrections

R_0 – resistance at the reference point without corrections

δR_{T1} – correction of any fixed point due to chemical impurity

δR_{T2} – correction of any fixed point due to hydrostatic pressure

δR_{T3} – correction of any fixed point due to perturbing heat exchanges

δR_{T4} – correction of any fixed point due to self-heating

δR_{T5} – correction of any fixed point due to AC/DC measurements

δR_{01} – correction of the reference point due to chemical impurity of water

- δR_{02} – correction of the reference point due to hydrostatic pressure
- δR_{03} – correction of the reference point due to perturbing heat exchanges
- δR_{04} – correction of the reference point due to self-heating
- δR_{05} – correction of the reference point due to AC/DC measurements
- δR_{06} – correction of the reference point due to SPRT internal insulation leakage.

Resistance of the thermometer at fixed point is given by equation

$$R_T = X_{Tc} \cdot R_{STc} = (X_T + \delta X_T)(R_S + \delta R_{S1} + \delta R_{S2}) \quad (2)$$

- X_{Tc} – reading of the bridge at the temperature of any fixed point after correction
- R_{STc} – resistance of the reference standard resistor after correction
- X_T – reading of the bridge at the temperature of any fixed point before correction
- R_S – resistance of the reference standard resistor before correction
($R_S = 25.000289 \Omega$)
- δX_T – “noise” of the bridge caused by random effects
- δR_{S1} – correction of reference standard instability (time drift)
- δR_{S2} – temperature correction of reference standard.

Input quantity connected with the resistance at the triple point of water:

$$R_0 = X_{0c} \cdot R_{S0c} = (X_0 + \delta X_0)(R_S + \delta R_{S3} + \delta R_{S4}) \quad (3)$$

- X_{0c} – reading of the bridge at the reference point after correction
 - R_{S0c} – resistance of the reference standard resistor after correction
 - X_0 – reading of the bridge at the reference point before correction
($X_0 = 1.0208720$)
 - R_S – resistance of the reference standard resistor before correction
($R_S = 25.000289 \Omega$)
 - δX_0 – “noise” of the bridge caused by random effects
 - δR_{S3} – correction of reference standard instability (time drift)
 - δR_{S4} – temperature correction of reference standard resistor.
- For the purpose of calculation of uncertainty all corrections are chosen to be equal to zero.

3. Combined uncertainty

Based on measurement equation the formula for combined uncertainty is given by Eq.4.

$$u_c^2(W) = c_1^2[u^2(X_T) + u^2(\delta X_T)] + c_2^2[u^2(X_0) + u^2(\delta X_0)] + c_3^2[u^2(R_S) + u^2(\delta R_{S1}) + u^2(\delta R_{S2})] + \\ + c_4^2[u^2(R_S) + u^2(\delta R_{S3}) + u^2(\delta R_{S4})] + 2c_3c_4u^2(R_S) \cdot r(R_S, R_S) + c_3^2 \sum_{i=1}^5 u^2(\delta R_{Ti}) + c_6^2 \sum_{i=1}^6 u^2(\delta R_{0i}) \quad (4)$$

where:

$$c_1 = \frac{\partial W}{\partial R_{Tc}} \frac{\partial R_T}{\partial R_{Tc}} \approx \frac{R_S}{X_0 R_S} = \frac{1}{X_0} \quad c_2 = \frac{\partial W}{\partial R_{0c}} \frac{\partial R_0}{\partial X_{0c}} \approx -\frac{X_T R_S}{X_0^2 R_S} = -\frac{X_T}{X_0^2}$$

$$c_3 = \frac{\partial W}{\partial R_{Tc}} \frac{\partial R_T}{\partial T_{STc}} \approx \frac{X_T}{X_0 R_S} \quad c_4 = \frac{\partial W}{\partial R_{0c}} \frac{\partial R_0}{\partial R_{S0c}} \approx -\frac{X_T X_0}{X_0^2 R_S} = -\frac{X_T}{X_0 R_S}$$

$$c_5 = \frac{\partial W}{\partial R_{Tc}} \approx \frac{1}{X_0 R_S} \quad c_6 = \frac{\partial W}{\partial R_{0c}} \approx -\frac{X_T}{X_0^2 R_S}$$

Taking into account that correlation coefficient $r(R_S, R_S)$ is equal 1 and $c_4 = -c_3$ we obtain simplified Eq.5

$$u_c^2(W) = c_1^2[u^2(X_T) + u^2(\delta X_T)] + c_2^2[u^2(X_0) + u^2(\delta X_0)] + c_3^2[u^2(\delta R_{S1}) + u^2(\delta R_{S2})] + \\ + c_4^2[u^2(\delta R_{S3}) + u^2(\delta R_{S4})] + c_3^2 \sum_{i=1}^5 u^2(\delta R_{Ti}) + c_6^2 \sum_{i=1}^6 u^2(\delta R_{0i}) \quad (5)$$

where:

- $u(X_T)$ – uncertainty of the F18 Bridge as stated in technical documentation
- $u(\delta X_T)$ – standard uncertainty calculated based on series of 10 measurements
- $u(X_0)$ – uncertainty of the F18 Bridge as stated in technical documentation
- $u(\delta X_0)$ – standard uncertainty calculated based on series of 20 measurements
- $u(\delta R_{S1})$ – standard uncertainty of standard resistor, caused by a drift in time (calculated as a half year instability)
- $u(\delta R_{S2})$ – standard uncertainty of standard resistor due to temperature variation (calculated based on technical data of F-18 Bridge for ± 5 K ambient temperature variation and thermostat stabilization coefficient equal 30)
- $u(\delta R_{S3})$ – standard uncertainty of standard resistor, caused by a drift in time (calculated as a one year instability)
- $u(\delta R_{S4})$ – standard uncertainty of standard resistor due to temperature variation (calculated based on technical data of F-18 Bridge for 5 K ambient temperature variation and thermostat stabilization coefficient equal 30)
- $u(\delta R_{T1})$ – standard uncertainty of achieving the equilibrium state at the triple point of hydrogen = 0.17 mK (value given by CCT/01-02 document) recalculated for uncertainty of thermometer resistance
- $u(\delta R_{T2})$ – standard uncertainty of hydrogen pressure measurement (due to imperfection of measurement of pressure)
- $u(\delta R_{T3})$ – standard uncertainty due to non-adiabatic conditions in cryostat based on series of 3 measurements
- $u(\delta R_{T4})$ – self-heating of resistance thermometer based on series of 10 measurements

- $u(\delta R_{T5})$ – standard uncertainty due to AC bridge excitation
- $u(\delta R_{01})$ – standard uncertainty of achieving the equilibrium state at the triple point of water 0.1 mK (value given by CCT/01-02 document), calculation performed with a use of $dR/dT = 0.1 \Omega/K$ resulting as $1E-5 \Omega$
- $u(\delta R_{02})$ – standard uncertainty of pressure measurement
- $u(\delta R_{03})$ – standard uncertainty due to temperature gradient inside the cell (evaluated in experimental way, by several trials of thermometer immersing based on series of 5 measurements)
- $u(\delta R_{04})$ – self-heating of resistance thermometer (current was 2 mA), based on series of 10 measurements (value estimated experimentally)
- $u(\delta R_{05})$ – standard uncertainty due to AC bridge excitation
- $u(\delta R_{06})$ – standard uncertainty due to wiring leakage.

4. Budget of uncertainties

The summarised uncertainties and their values, applied distribution functions plus sensitivity coefficients and contributions of each component are presented in Table 1.

Tab. 1. Summarized uncertainties and their probability distribution functions (for triple point of hydrogen and triple point of water)

Quantity	Estimate		Standard uncertainty		Probability distrib.	Deg free.	Sensitivity coefficient		Contribution
X_T	0.0011901		1.16E-07		rectang.		0.9795547		1.131E-07
δX_T	0		5.12E-09		t-Student	9	0.9795547		5.015E-09
X_0	1.020872		1.16E-07		rectang.		-0.0011419		-1.319E-10
δX_0	0		2.16E-09		t-Student	19	-0.0011419		-2.467E-12
δR_{S1}	0	Ω	1.44E-05	Ω	rectang.		4.663E-05	1/ Ω	6.730E-10
δR_{S2}	0	Ω	1.64E-07	Ω	rectang.		4.663E-05	1/ Ω	7.628E-12
δR_{S3}	0	Ω	2.89E-05	Ω	rectang.		-4.663E-05	1/ Ω	-1.346E-09
δR_{S4}	0	Ω	1.64E-07	Ω	rectang.		-4.663E-05	1/ Ω	-7.628E-12
δR_{T1}	0	Ω	1.09E-06	Ω	normal		0.03918174	1/ Ω	4.263E-08
δR_{T2}	0	Ω	4.62E-08	Ω	rectang.		0.03918174	1/ Ω	1.809E-09
δR_{T3}	0	Ω	3.36E-09	Ω	t-Student	2	0.03918174	1/ Ω	1.316E-10
δR_{T4}	0	Ω	0.00E+00	Ω	t-Student	9	0.03918174	1/ Ω	0
δR_{T5}	0	Ω	2.17E-06	Ω	rectang.		0.03918174	1/ Ω	8.483E-08
δR_{01}	0	Ω	1.00E-05	Ω	normal		-4.568E-05	1/ Ω	-4.568E-10
δR_{02}	0	Ω	4.04E-07	Ω	rectang.		-4.568E-05	1/ Ω	-1.846E-11
δR_{03}	0	Ω	5.00E-06	Ω	t-Student	4	-4.568E-05	1/ Ω	-2.283E-10
δR_{04}	0	Ω	1.10E-06	Ω	t-Student	9	-4.568E-05	1/ Ω	-5.024E-11
δR_{05}	0	Ω	2.89E-06	Ω	rectang.		-4.568E-05	1/ Ω	-1.318E-10
δR_{06}	0	Ω	7.22E-07	Ω	rectang.		-4.568E-05	1/ Ω	-3.296E-11
W	0.0011657						u_c	–	1.478E-7

5. Calculation of expanded uncertainty

5.1. Conventional method

For conventional method coverage factor is assumed as $k = 2$ for confidence level $p = 0.95$

5.2. Approximate method

The following value of coverage factor is assumed, depending on ratio, which in this case was defined as:

$$r_u = \frac{|u_i(y)|}{\sqrt{u_c^2(y) - u_i^2(y)}}$$

where: u_c – geometric sum of all contributing components

u_i – the largest contributing component of type B.

There 3 cases are considered as:

a) $k=k_N$ for $0 < r_u \leq 1$ (normal distribution) $k_N = 1.96$ ($p=95\%$)

b) $k=k_T$ for $1 < r_u \leq 10$ (trapezoidal distribution)

$$\text{where: } k_T = \frac{\sqrt{3}}{\sqrt{r_u^2 + 1}} (1 + r_u - 2\sqrt{r_u(1-p)})$$

c) $k=k_p$ for $r_u > 10$ (uniform distribution) $k_p = \sqrt{3} p$.

Details are described in the article titled: “A method of approximation of the coverage factor in calibration” by P. Fotowicz, accepted for publication in the journal *Measurement* 11 December 2003

5.3 Convolution method

Block diagram of expanded uncertainty calculation based on convolution method is presented in another paper in this book. In this case all contributing components are taken into account and k-factor is no needed to calculate expanded uncertainty. In such a type of consideration the term of expanded uncertainty should be replaced by term: coverage interval at p level of confidence.

Example of one calculation of coverage interval based on convolution method in the form of graphical representation of convolutions of probability density functions of contributing components, and final convolution and its density function are presented in Figure 1.

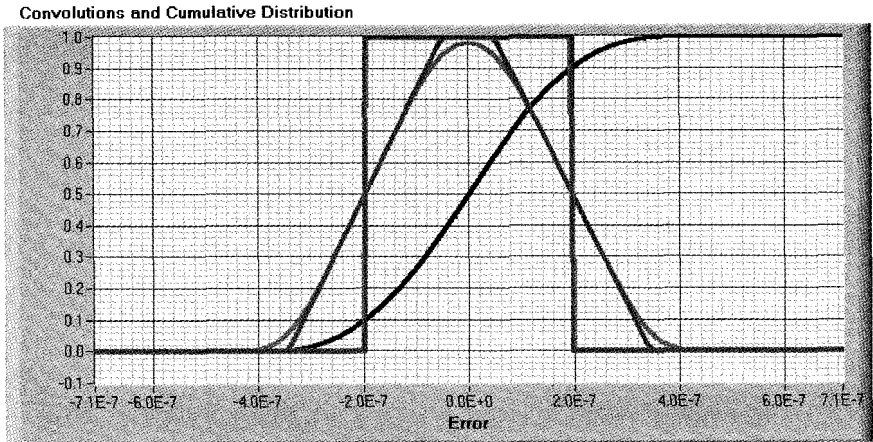


Figure 1. Example of screen intermediate calculations of coverage interval: convolutions of probability distribution functions, and cumulative distribution.

6. Comparison of results

Fixed point	W	Method I		Method II		Method III	
		k	U	k	U	k	U
Hydrogen	0.001165768	2.00	2.96 E-07	1.90	2.80 E-07	1.90	2.80 E-07
17.0 K	0.00223809	2.00	1.92 E-06	1.96	1.87 E-06	1.96	1.87 E-06
20.3 K	0.00419230	2.00	3.02 E-06	1.96	2.96 E-06	1.96	2.96 E-06
Neon	0.008311424	2.00	5.19 E-07	1.96	5.08 E-07	1.99	5.17 E-07
Oxygen	0.08991000	2.00	1.59 E-06	1.96	1.56 E-06	1.97	1.57 E-06
Argon	0.21150800	2.00	1.62 E-06	1.96	1.59 E-06	1.96	1.58 E-06
Mercury	0.84415480	2.00	2.48 E-06	1.90	2.36 E-06	1.92	2.38 E-06

7. Conclusion

Although the results of calculation performed by method II and method III are very similar the advantage of method III is clear; no subjective and rather sophisticated deliberations are needed.

References

1. P. Fotowicz, Method for calculating the coverage factor in calibration. *OIML Bulletin*, vol. XLIII, no 4, (October 2002)
2. M. J. Korczynski, P. Fotowicz, A. Hetman, Calculation of Expanded Uncertainty in Calibration of RTD Sensors at Low Temperature, *UNCERN 2003, Oxford*, (April 2003).

A NEW APPROACH TO THE PRESENTATION OF THE RESULT MEASUREMENTS IN VIRTUAL INSTRUMENTS*

M. J. KORCZYNSKI AND A. HETMAN

*Technical University of Lodz,
Institute of Theoretical Electrotechnics, Metrology and Material
18/22 Stefanowskiego Street, 90-924 Lodz, Poland
E-mail: jerzykor@p.lodz.p*

The main purpose of the paper is to present how easily and clearly the result of measurements can be accompanied by coverage interval calculated at any desired confidence level in Virtual Instrument, if an appropriate procedure is implemented in instrument, or if instrument is equipped with additional processor, which handles measured data as a series of values, calculates coverage interval, which covers the true value at a certain level of confidence.

1. The challenge

The new way of data presentation in virtual or any other instruments proposed here is: each displayed result of measurements is accompanied by a parameter characterizing its dispersion in a way attributable to the measured value with a certain degree of confidence [2, 4]. The desired confidence level is declared by a person performing the measurements.

Such a presentation can be implemented directly in virtual instruments as only a small amount of software, presented here, needs to be added to the overall software. A specialized processor can play the role of this software.

This method fully describes the coverage region, which covers the true measured value.

2. The method of uncertainty calculation

Most contemporary measurement instruments are digital, thus block diagram representing their structure can be recognized. The PC station presented in Fig 1 is typical for virtual instrument. In other digital instruments there are elements which play similar role from functional point of view.

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

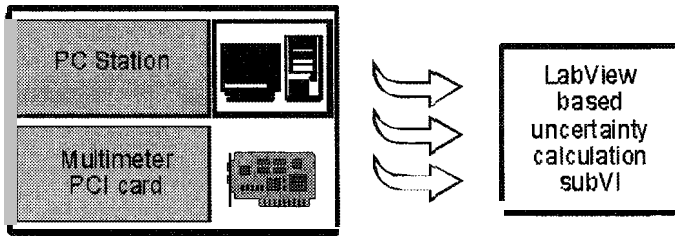


Figure 1. Block diagram of a virtual instrument equipped with Multimeter PCI card.

The algorithm of developed software is based on idea presented in block diagram depicted in Fig. 2.

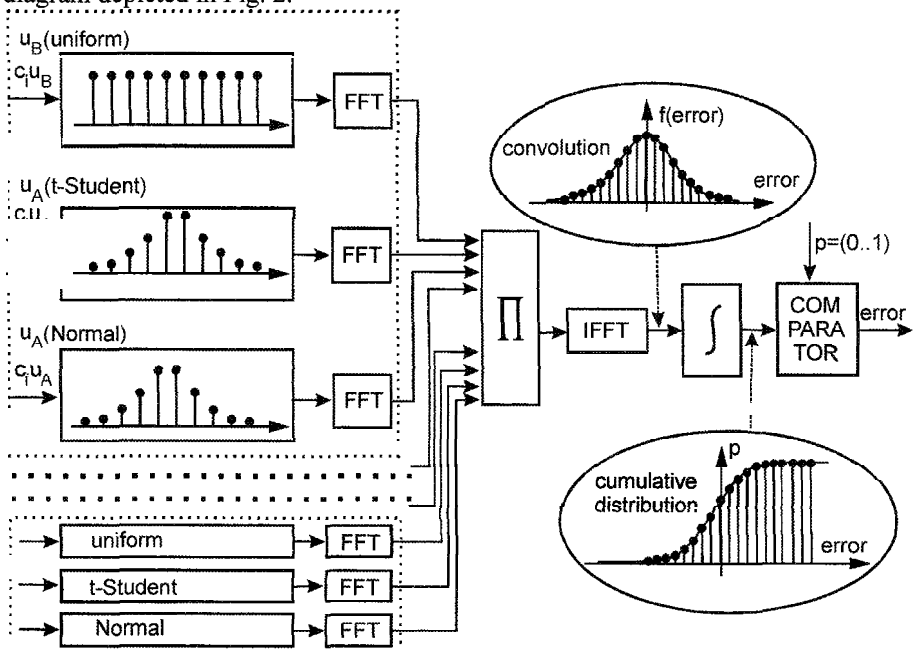


Figure 2. Block diagram of developed software for calculation coverage interval with a p – level of confidence, FFT – Fast Fourier Transform, Π – Product, IFFT – Inverse Fast Fourier Transform, c_i – sensitivity factors, u_A – uncertainties of type A, u_B – uncertainties of type B

Parameters of any A/D input card, or any sensor, any transducer, or converter stated in the technical specification are the input data for software. The desired level of confidence for coverage interval is declared by the experimenter, performing measurements. Additionally a number of individual sample measurements, i.e. a number of repeated readings for internal calculation of mean, and parameters to calculate type A uncertainties in the elaborated software, is also declared by experimenter. The parameters and the type of

probability distribution functions, describing the digital and analogue accuracy of the instrument component, must be supplied by hardware producer. The probability distribution function for repeated individual readings, are assumed to be randomly distributed according to t-Student distribution.

3. The results

The three examples are the images of virtual user panel of the virtual instrument elaborated using LabView (NI license)

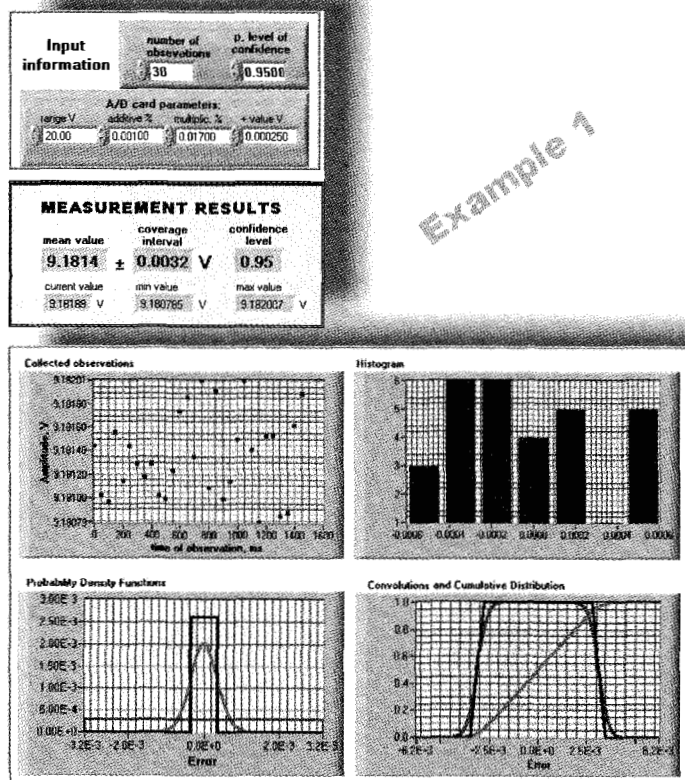


Figure 3. Human-Machine Interface of Virtual Voltmeter with 30 declared observations and results of type A and type B uncertainties calculation.

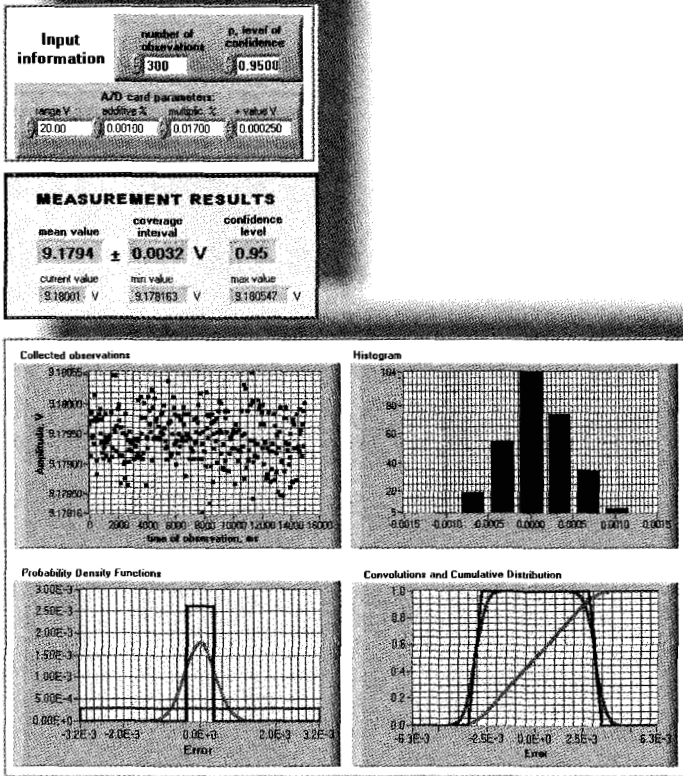


Figure 4. Human-Machine Interface of Virtual Voltmeter with 300 declared observations and results of type A and type B uncertainties calculation.

4. Conclusions

The new way of data presentation will not leave any doubts to any experimenter, where the true value is located.

At today's state of technology it is not necessary to avoid quoting the uncertainty of measurement results with LCD display.

All digital instruments can be equipped with a programme or specialized processor, which calculates the uncertainty at any required level of confidence.

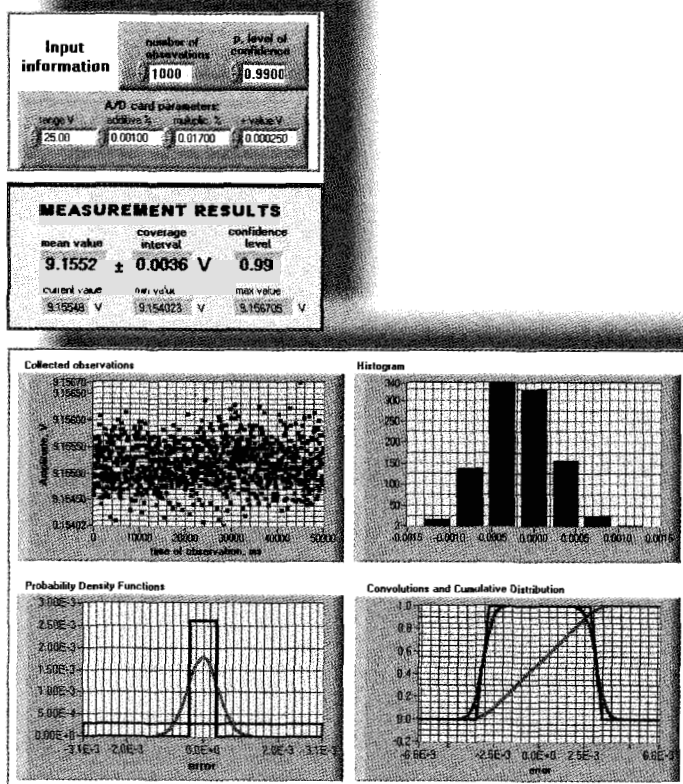


Figure 5. Human-Machine Interface of Virtual Voltmeter with 1000 declared observations and results of type A and type B uncertainties calculation.

References

1. K. Gniotek, Creative Measurement in View of the Scope of Metrology, *Measurement – Journal of the IMEKO*, vol. 20, No. 4, 259, (1997)
2. P. Fotowicz, Method for calculating the coverage factor in calibration. *OIML Bulletin*, vol. XLIII, no 4, (October 2002)
3. S. Kubisa, On Probability Distribution of the Instrument Error, *Metrology and Measurement Systems*, vol. VIII, No 4, 303, (2001).
4. EAL-R2 Expression of the Uncertainty of Measurement in calibration, Edition 1, (April 1997).

A HYBRID METHOD FOR ℓ_1 APPROXIMATION ^a

D. LEI, J. C. MASON

*School of Computing and Engineering,
 University of Huddersfield, Huddersfield, HD1 3DH, UK
 Email: d.lei@hud.ac.uk
 j.c.mason@hud.ac.uk*

Linear programming (LP) techniques and interior-point methods (IPMs) have been used to solve ℓ_1 approximation problems. The advantage of the IPMs is that they can reach the vicinity of the optimum very quickly regardless of the size of the problems, but numerical difficulties arise when the current solution is approaching optimal. On the other hand, the number of iterations needed for an ℓ_1 approximation problem by LP techniques is proportional to the dimension of a problem. However these LP methods are finite algorithms, and do not have the problems that the IPMs endure. It would be beneficial to combine the merits of both methods to achieve computational efficiency and accuracy. In this paper, we propose an algorithm which applies the IPMs to get a near best solution and fine-tune the results by one of the simplex methods. The savings in terms of numbers of iterations can be substantial.

1 Introduction

A linear discrete ℓ_1 approximation problem can be stated as: Given a set of discrete data points $\{(x_i, y_i)\}_{i=1}^m$, determining a set of parameters $c \in \mathbb{R}^n$, which can minimize the following term

$$\min_c \|y - Ac\|_1 = \sum_{i=1}^m \left| y_i - \sum_{j=1}^n a_{i,j} c_j \right| = \sum_{i=1}^m |r_i|,$$

where A is the observation matrix, $A \in \mathbb{R}^{m \times n}$, $m \geq n$ and $r_i = y_i - \sum_{j=1}^n a_{i,j} c_j$.

The ℓ_1 approximation problem has been successfully solved by linear programming techniques, the Barrodale and Roberts method ¹ is the usual algorithm of preference. However, it has long been known that the number of iterations required by a simplex-based method tends to be linear in the problem size, and in the worst case, it could grow exponentially with the problem.

Several authors have tempted to solve the ℓ_1 approximation problem by the interior-point methods. For example, Ruzinsky and Olsen proposed a simple iterative algorithm for ℓ_1 approximation problem ² on a variant of Karmarkar's

^a"Work partially funded under EU Project SoftTools_MetroNet Contract N. G6RT-CT-2001-05061"

linear programming algorithm³. One of the salient features of the IPMs-based approach is that the objective function decreases rapidly in the first few iterations, but the convergence process slows down when the current solution is approaching the optimum. Solving ℓ_1 approximation problem by IPMs could be computationally expensive when exact solutions are required. In this paper, we shall introduce a hybrid method for ℓ_1 approximation. It is a combination of the central and boundary search schemes, which starts with an interior-point method and switches to a simplex-based approach at an appropriate point. The effectiveness of the proposed method can be shown by various experimental results.

2 The hybrid method

The hybrid method starts with an interior-point method – Ruzinsky and Olsen’s algorithm². It considers the ℓ_1 fitting problem in the dual form of

$$\begin{aligned} & \text{maximize} && u^T y - v^T y \\ & \text{subject to} && A^T u - A^T v = 0 \\ & && u + v = e \end{aligned} \tag{1}$$

where e is a unit vector with all entries equal to 1.

The major computational work in each iteration of an IPMs involves solving a weighted linear least squares approximation problem. Ruzinsky and Olsen investigated the characterization of an ℓ_1 solution and proposed the following weighting scheme:

$$w_i = \frac{u_i^2 v_i^2}{u_i^2 + v_i^2} \quad \text{and} \quad W = \text{diag}(w). \tag{2}$$

where W is the weighting matrix.

The algorithm terminates when

$$1 - \frac{\sum_i |\zeta_i|}{\sum_i |r_i|} < \epsilon. \tag{3}$$

where $\zeta_i = \frac{u_i - v_i}{u_i^2 + v_i^2} r_i$ and ϵ is a predetermined tolerance value.

The author claimed that such a stopping rule is well behaved especially in the presence of roundoff errors. In our numerical experiences, it terminates the computation when (3) is satisfied for some predetermined ϵ . But one important observation we have noticed is that even if the Ruzinsky and Olsen’s algorithm

converges to the required accuracy, which does not imply the actual solution is within the stated accuracy.

Having obtained a near best solution, we apply Barrodale and Roberts algorithm¹ to associate this interior point with a vertex. Assume that readers have sound knowledge of the Barrodale and Roberts algorithm. The hybrid method can be outlined as follows:

1. Set a rough accuracy tolerance value of $\epsilon = 0.1$, apply Ruzinsky and Olsen's algorithm to get a near best solution u and v .
2. Compute the corresponding coefficient parameters c by solving the weighted least squares problem $WAc = Wy$, where W is constructed by (2). Obtain the residual vector by $r = y - Ac$.
3. Fit the residuals r by the Barrodale and Roberts algorithm to obtain a correcting approximating function with correcting coefficients δc .
4. Compute the optimal solution $c^* = c + \delta c$.

The mathematical justification of the hybrid method can be derived as

$$\begin{aligned}\|y - Ac^*\|_1 &= \|y - A(c + \delta c)\|_1 \\ &= \|r - A\delta c\|_1.\end{aligned}$$

Numerical experiments have been carried out to test the effectiveness of the proposed method. Some of the results are rather striking, the number of iterations required by the Barrodale and Roberts algorithm decrease dramatically by using the results from the IPMs.

3 Conclusion

The proposed hybrid method has incorporated the IPMs and LP techniques, which builds a bridge between these two important optimization techniques. This pilot study indicates a possibility of solving large ℓ_1 approximation problems by taking advantages of both methods.

References

1. I. Barrodale and F. D. K. Roberts. An improved algorithm for discrete ℓ_1 linear approximation. *SIAM Journal of Numerical Analysis*, 10(5):839–848, 1973.

2. S. A. Ruzinsky and E. T. Olsen. ℓ_1 and ℓ_∞ minimization via a variant of Karmarkar's algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing*. 37(2):245–253, 1989.
3. R. J. Vanderbei, M. S. Meketon and B. A. Freedman. A modification of Karmarkar's linear programming algorithm. *Algorithmica*. 1(4):395–407, 1986.

INTERPOLATION EQUATIONS FOR INDUSTRIAL PLATINUM RESISTANCE THERMOMETERS*

P. MARCARINO, P.P.M. STEUR AND A. MERLONE

CNR Istituto di Metrologia "Gustavo Colonnetti", Torino, Italy

The Callendar – Van Dusen (CVD) equations of the IEC-751 for the Industrial Platinum Resistance Thermometers (IPRTs) do not exactly follow the International Temperature Scale of 1990 (ITS-90), within the required accuracy. Indeed, there is a demand for calibrations of IPRTs to uncertainties of around 10 mK for temperatures from 0 to 250 °C and better than 0.1 K between 250 °C and about 500 °C, while the use of the CVD equations do not allow an interpolation uncertainty better than ± 0.2 K over a range larger than 0 - 250 °C. To solve this problem, two new reference equations, one below 0 °C and the other above 0 °C, are proposed to be used instead of the CVD equations as reference for the IPRTs. These equations will be of a higher order than the CVD equations and their lower order coefficients are equal to the constants A , B and C of the CVD. The use of these new reference equations allows an interpolation accuracy at the millikelvin level, limiting the number of calibration points in every range to five.

1. Introduction

Before 1968, the Callendar-Van Dusen (CVD) equations
 for $t < 0$ °C

$$\frac{R_t}{R_0} = \left[1 + A \cdot t + B \cdot t^2 + C(t - 100 \text{ °C})t^3 \right] \quad (1)$$

for $t > 0$ °C

$$\frac{R_t}{R_0} = \left[1 + A \cdot t + B \cdot t^2 \right] \quad (2)$$

described the resistance-temperature relation of both standard and industrial platinum resistance thermometers, indicated respectively with SPRTs and IPRTs [1, 2]. Since 1968, the CVD equations were used only for the IPRTs and their agreement with the SPRTs realizing the temperature scale was to be verified [3, 4, 5, 6].

From 1990, with the introduction of the ITS-90 [7], temperature t_{90} was defined by means of a reference function, a logarithmic polynomial equation of 12th order below 0 °C and of 9th order above 0 °C, and a quadratic linear deviation function, the constants of the latter being determined by calibration at

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

the fixed points. The differences between the ITS-90 and the CVD equations are shown in Figure 1.

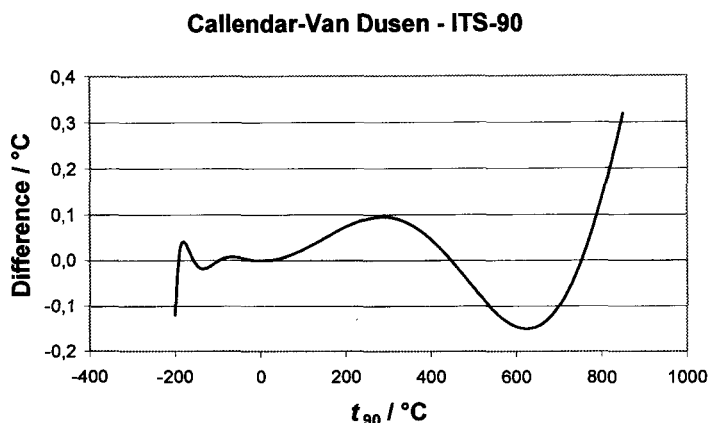


Figure 1. Differences between the CVD equations and the ITS-90 in the temperature range from $-200\text{ }^\circ\text{C}$ to $850\text{ }^\circ\text{C}$.

Many laboratories are using platinum resistance thermometers that do not satisfy the requirements of the ITS-90 as working standard for temperature calibrations. The working standard thermometer needs to be inexpensive and to exhibit very good stability, but only in the operating range of its bath or furnace. Thermometer calibrations can be carried out either directly by comparison with the standard reference thermometer (for best accuracy) or, as for routine calibrations, by comparison with the working standard only.

The calibration accuracy depends on the stability of the working standard between subsequent calibrations, and also on the capability of the interpolating equation to approximate the ITS-90. As is evident from Figure 1, the CVD equations are completely inadequate at the level of $0.01\text{ }^\circ\text{C}$.

In the mean time, calibration service laboratories have improved their capabilities considerably and some have been recognized at the level of $0.03\text{ }^\circ\text{C}$ in a limited range, and it is to be expected that their capabilities will continue to improve. Evidently an alternative, adequate, interpolation equation is required.

2. The proposed reference functions

The ITS-90 reference function has been fitted with two functions, one for $t < 0\text{ }^\circ\text{C}$, between $-200\text{ }^\circ\text{C}$ and $0\text{ }^\circ\text{C}$, and one for $t > 0\text{ }^\circ\text{C}$, between $0\text{ }^\circ\text{C}$ and

850 °C, maintaining for the lower orders the form of the CVD equations. They are of the form:

for $t < 0$ °C

$$\frac{R_t}{R_0} = \left[1 + A \cdot t + B \cdot t^2 + C(t - 100 \text{ °C})t^3 + \sum_5^{12} D_i \cdot t^i \right] \quad (3)$$

for $t > 0$ °C

$$\frac{R_t}{R_0} = \left[1 + A \cdot t + B \cdot t^2 + \sum_3^{11} C_i \cdot t^i \right] \quad (4)$$

where the coefficients A , B and C are those of the CVD equations and the coefficients D_i and C_i have been calculated in order to obtain fits with residuals less than 1 mK over *all* of the range. All coefficient values are reported in Table I.

Table I The coefficients for the two proposed reference functions, eqs. (3) and (4).

A	0.0039083	A	0.0039083
B	-5.775E-7	B	-5.775E-7
C	-4.183E-12	C_3	-2.48696E-10
D_5	1.14142E-12	C_4	2.62329E-12
D_6	5.24730E-14	C_5	-1.41357E-14
D_7	1.04763E-15	C_6	4.78094E-17
D_8	1.17388E-17	C_7	-1.04074E-19
D_9	7.92701E-20	C_8	1.44730E-22
D_{10}	3.21180E-22	C_9	-1.24395E-25
D_{11}	7.21836E-25	C_{10}	6.03317E-29
D_{12}	6.94960E-28	C_{11}	-1.26471E-32

3. Application of the reference functions to actual calibration data

The reference functions obtained above can be fitted to the IPRTs calibration data in two ways: (a) by calculating the A , B and C coefficient values of eqs. (3) and (4), where all coefficients D_i and C_i are constants, in the same way as for the CVD equations; (b) or by using a 2nd order deviation function with respect to eqs. (3) and (4) as reference functions, in the same way as for thermocouples.

The described interpolating methods have been applied to the same calibration data from SIT laboratories (Italian Calibrations Services) as used in

[6]. Figure 2 shows the results of a straightforward application of the CVD

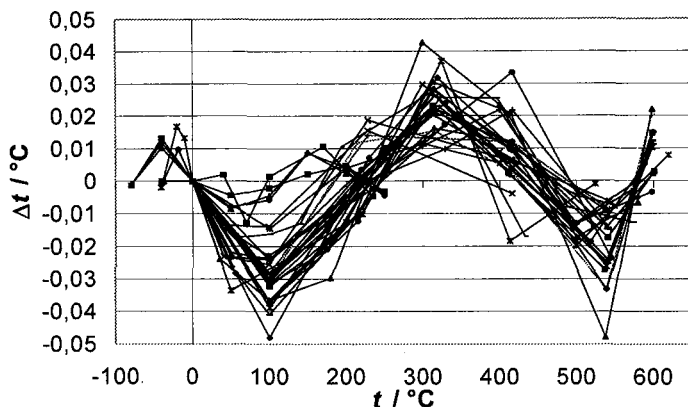


Figure 2 The residuals of the straightforward application of the CVD equations to the calibration data of 38 working standard thermometers of SIT laboratories.

equations, while Figure 3 shows the residuals of the application of the *new reference functions*.

The improvement in residuals from Figure 2 to Figure 3 is more than evident, while the remaining scatter mirrors the calibration capabilities of the single laboratories. Therefore the adoption of the reference functions proposed in this work satisfies the increasing demand for calibration accuracy of IPRTs to the Calibration Services and will leave ample space for improvement in measurement capability.

4. Conclusions

Two new reference functions, one below 0 °C and the other above 0 °C, are proposed as substitutes for the CVD equations as standard functions for the IPRTs. The new reference functions have the following advantages:

- 1) maintain for the lower order coefficients the same constants A , B and C of the CVD;
- 2) the higher order coefficients are constant for all IPRTs and allow an agreement with the ITS-90 well within 1 mK;
- 3) only the constants A , B and C are to be calculated from the calibration points, that can be then limited to five over all ranges.

5. References

1. Comptes Rendus, Onzième Conf. Gén. Des Poids et Mesures (Paris, 1960). J. Research Natl. Bur. Standards, 1961, **65A**, 139.
2. IEC – Industrial Platinum Resistance Thermometer Sensors, *IEC Standard Publication 751* (Bureau Central de la Commission Electrotechnique Internationale, Geneva), Amendment 2, 1995.
3. Connolly J.J., In *Temperature: Its Measurement and Control in Science and Industry*, Vol. 5 (Edited by J. F. Schooley), New York, American Institute of Physics, 1982, 815-817.
4. Actis A. and Crovini L., In *Temperature: Its Measurement and Control in Science and Industry*, Vol. 5 (Edited by J. F. Schooley), New York, American Institute of Physics, 1982, 819-827.
5. Crovini L., Actis A., Coggiola G., Mangano A., In *Temperature: Its Measurement and Control in Science and Industry*, Vol. 6 (Edited by J.F. Schooley), New York, American Institute of Physics, 1992, 1077-1082.
6. P.Marcarino, P.P.M.Steur, G.Bongiovanni, B.Cavigioli, 2002: "ITS-90 Approximation by means of non-standard platinum resistance thermometers", Proc. TempMeko2001 (ed. B.Fellmuth, J.Seidel, G.Scholz), Berlin (Germany), 85-90.
7. The International Temperature Scale of 1990, *Metrologia*, 1990, **27**, 3-10.

FROM THE FIXED POINTS CALIBRATION TO THE CERTIFICATE: A COMPLETELY AUTOMATED TEMPERATURE LABORATORY*

A.MERLONE, P.MARCARINO, P.P.M. STEUR, R. DEMATTEIS

*Istituto di Metrologia “G. Colonnetti”, IMGC-CNR, Strada delle Cacce 73, 10135
Torino, Italy*

E-mail: a.merlone@imgc.cnr.it

At the Intermediate Temperature Laboratory of the Istituto di Metrologia “G. Colonnetti” (IMGC) all operations concerning the calibration of standard platinum resistance thermometers (SPRTs) at the ITS-90 fixed points are now completely automated. The software is based upon Visual Basic® modules, forms, macros, I/O connection and dialogs with the main Microsoft® Office® application. These modules and other “.exe” files are also used for research purposes as they can be separately used as virtual consoles for data acquisition, pre-processing and post-processing. Several blocks call on each other, starting from the data acquisition to the certificate printout: those blocks can be used as useful tools for accredited calibration laboratories. Statistics, data input evaluation, cross controls, ITS-90 requirements and equations, fixed points and thermometer databases are in continuous dialog, till the end of the calibration procedure. All the data, both from the calibration and the research activities is weekly saved automatically.

1. Introduction

The Intermediate Temperature Laboratory of the Istituto di Metrologia “G. Colonnetti” (IMGC) is responsible for the ITS-90 [1] maintenance and the calibration of SPRTs at the fixed points [2]. The laboratory performs SPRT calibrations for research activities, for the contact thermometry laboratory of the IMGC calibration service and for several calibration laboratories of SIT (Italian Calibrations Services). All the calibration procedures, from the resistance measurements to the certificate issue, are now completely automated.

The programs are continuously running, twenty four hours a day, in order to keep under strict control all the single parts involved in the whole system. During those times free from calibration or research activities several noise analyses, instruments stability measurements and I/O tests are performed.

Each single fixed point calibration takes into account all the requested procedures such as the temperature stability, self-heating and the standard deviation of a pre-set number of readings. The whole measurement sequence is automated. The operator just pushes a virtual button and all the data start to be

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

acquired, evaluated, displayed and saved. The calibration procedures are now easier and quicker and the whole system is flexible to several kind of requirements, from research to industrial PRT calibration. The software has been completely developed at IMGC. Several blocks can be used in future as validated software by metrological laboratories for accredited temperature measurements.

2. The procedure

As the SPRT is inserted in the thermometer well of the fixed point cell, the complete process for the calibration is started. Many virtual consoles take part in the data acquisition, from the instruments connected to the SPRT. The implementation of object-oriented systems has led to the realization of those ActiveX modules specialized in the management of instrumentation, that show the same interface to the outside world. These consoles allow a complete control of all the instruments involved in the calibrations. The user-interfaces that have been developed reflect the potential of the instruments, with useful functions for the user, directly available through list-boxes, check-boxes and command-buttons, as well as the potential of the software, that allow the totally automated operation of any sequence. The instruments are all connected to PCs and all the I/O information fluxes are recorded and controlled.

The calibration uncertainty is automatically evaluated taking into account the stability of the thermometer at the triple point of water before and after each fixed point determination, the fixed point reproducibility, ITS-90 specifications and all the other A-type and B-type uncertainty components; several subroutines are involved in those processes. A preliminary statistical analysis of the data acquired is evaluated in real time by the virtual consoles that drive the acquisition instruments. Any malfunctioning is immediately evaluated and the operator can decide whether to repeat the calibration or to stop it. Next, the series of fixed point calibrations for the required ITS-90 sub-range are stored in the data-acquisition database that is used both for emission of the final certificate and as a complete record of the “history” of all thermometers.

Several Excel[®] macros are automatically recalled from the procedure and are used to evaluate the calibration constants from the ITS-90 requirements, considering the temperature range and the related temperature fixed points involved in the specific thermometer calibration.

The hydrostatic temperature correction is automatically taken into account, considering the cell used and the thermometer dimensions. The information regarding the fixed point cells are those evaluated during the CCT Key

Comparisons and used for the CMC [2]. They are automatically recalled from the specific database.

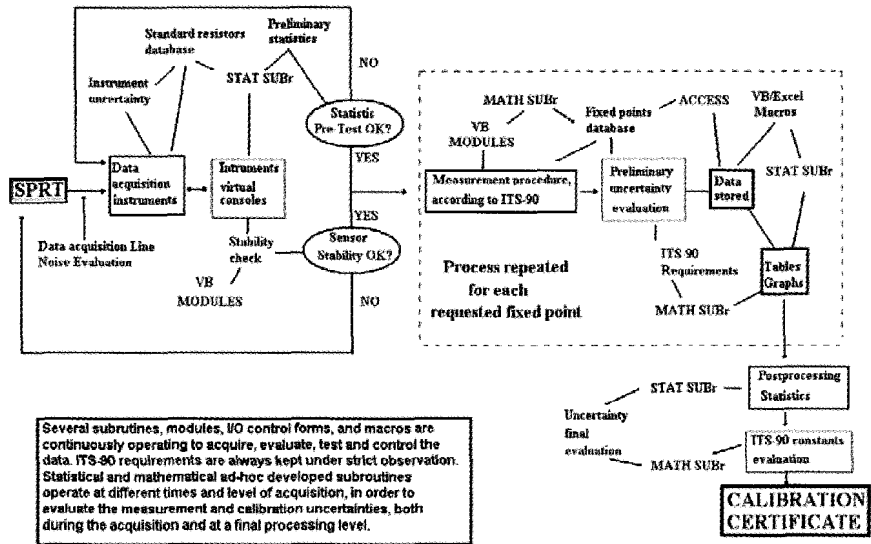


Figure 1. Schematic design of the blocks involved in the whole calibration process

The processed data is then sent to a Word[®] document that uses the Institute's model for the calibration certificate. This model is then automatically compiled with all the data requested, then printed. This will then be the calibration certificate, ready to be signed and sent with the SPRT. The complete procedures, the I/O devices and sequences, the data acquisition modules and the databases are continuously controlled and cross-checked. The whole procedure is a part of the laboratory quality system, as well as all the instrument and temperature standard procedures. The developed software involves the main Microsoft[®] Office[®] applications and dedicated data-acquisition modules [3], written in Visual Basic[®]. Those modules are also used for research purposes as they can be separately used as virtual consoles or acquisition modules. They represent useful tools for a metrological laboratory in the field of temperature data acquisition and can be validated as single parts for all the requirements of the ITS-90 specifications.

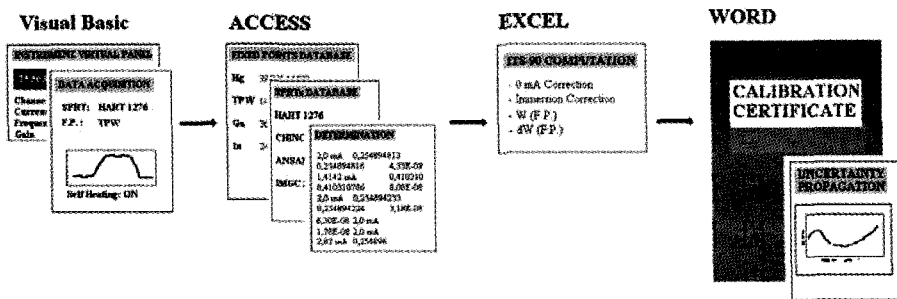


Figure 2. System logical diagram.

The structure that has been realized is therefore made up of a number of servers specialized in managing instruments (remote consoles) and some client programs specialized in the realization of specific functions. Once the modules for communication with the instruments have been created, the management programs become users of the service supplied by these same modules and represent in synthesis the requirements of the user; in this aspect they can be developed with those instruments that are most apt for their use.

The Word[®] document contains the evaluated calibrations constants, each fixed point W value measured, dW/dT tables (see ITS-90 specifications) and the associated uncertainty value and curves. All data, both from the calibration and the research activities is weekly saved automatically on two different backup directories on two separate computers.

References

1. H. Preston-Thomas, *The International Temperature Scale of 1990 (ITS-90)*, Metrologia 27, p. 3-10 (1990).
2. P. Marcarino, P.P.M. Steur, R. Dematteis, *Realization at IMGC of the ITS-90 Fixed Points from the Argon Triple Point upwards*, proceedings of "The 8th Symposium on Temperature: its measurement and control in Science and Industry", October 21-24, 2002, Chicago, Illinois, USA, p- 65 - 70.
3. B.Cavigioli, P.Marcarino, A.Merlone, P.P.M. Steur, *ActiveX-tools for data acquisition with automatic ASL-F18 and MI-6010, Advanced Mathematical and Computational Tools in Metrology*, vol.5, Series on Advances in Mathematics for Applied Sciences vol. 57, World Scientific, Singapore, 2001, p. 47-54.

A NEW OFF-LINE GAIN STABILISATION METHOD APPLIED TO ALPHA-PARTICLE SPECTROMETRY

S. POMMÉ AND G. SIBBENS

*EC-JRC-IRMM,
Institute for Reference Materials and Measurements
Retieseweg 111,
Geel, B-2440, Belgium
E-mail: stefaan.pomme@cec.eu.int*

A new off-line gain stabilisation method is applied to high-resolution alpha-particle spectrometry of ^{235}U . The software package SHIFTER automatically identifies and quantifies gain shift for intermediate spectra or even individual data in list mode files. By reversing the gain shift before combining all data into one sum spectrum, one can optimise the overall resolution. This automatic procedure is very useful with high-resolution spectrometry at low count rate, as a compensation for gain drift during the long measurement.

1. Introduction

A limitation of alpha-particle spectrometry arises from spectral interference caused by the presence of nuclides with near identical alpha-emissions, which appear as unresolved or partially resolved multiplets in the measured energy spectrum. As the optimum resolution requires thin radioactive sources, small detectors and relatively high source-detector distances, one needs long measurement times to obtain the required statistical accuracy. Hence, electronic gain shift with time can significantly influence the overall resolution.

In this work, we apply a new off-line gain shift correction method [1] to a series of 150 sequential high-resolution alpha-particle energy spectra (and the corresponding list mode data files), taken over a period of six months. The ‘Stieltjes integral’ of the measured spectra with respect to a reference spectrum is used as an indicator for gain instability. ‘Exponentially moving averages’ of the latter show the gain shift as a function of time. With this information, the data are relocated on a spectrum-by-spectrum or a point-by-point basis.

Advantages of this method are the flexibility involved with the off-line analysis, allowing for optimisation by iterative processes, and the optimal use of scarce information inherent to measurements at low count rate. It is an alternative to other correction techniques like e.g. an electronic gain stabiliser [2], which requires a distinct energy peak with sufficient count rate, or a manual

shift of individual spectra by a discrete number of channels, based on a visual check of certain peak positions.

2. Gain stabilisation method

In the case of alpha-particle spectrometry, most radionuclides are long-lived and the spectral shape can be assumed constant throughout the measurement campaign. The relevant alpha peaks for ^{235}U are situated between 4.15 and 4.6 MeV, corresponding to only 10% of the energy range of the measured spectrum. Hence, a modest gain instability results in a shift of the region of interest by practically the same amount of channels. The human eye recognises gain shifts by the displacement of spectral features (peaks, valleys) with respect to their position in a previous spectrum, acting as a reference spectrum. Off-line gain stabilisation implies quantification and counteraction of the shift for each measured spectrum (or data point).

Mathematically one can align two spectra e.g. by maximising their mutual convolution. A more convenient way of estimating the shift is based on the Stieltjes integral of a spectrum, $f_{\text{shifted}}(x)$, with respect to the reference spectrum, $f(x)$:

$$\text{shift} \approx \frac{\int_a^b f_{\text{shifted}}(x) \cdot f'(x) dx - \frac{1}{2} [f(b)^2 - f(a)^2]}{-\int_a^b f'(x)^2 dx \cdot \int_a^b f_{\text{shifted}}(x) dx \bigg/ \int_a^b f(x) dx} \quad (1)$$

The second term in the numerator is negligible if the integration is performed between two points where the count rate is extremely low, i.e. $f(a) = f(b) = 0$. The alpha particle spectra are taken with high resolution (8K channels), so that the spectrum $f(x)$ is a continuous and slowly varying function from one channel to the next. The Stieltjes integral combines the number of counts in each channel with the first derivative in the corresponding channel of the reference spectrum. The predominance of counts in channels with a negative $f'(x)$ value is interpreted as a drift in the forward direction, and vice versa. The denominator acts as a normalisation factor, based on the consideration that $f'(x) = f_{\text{shifted}}(x) - f(x)$, if the shift is exactly one channel, i.e. $f_{\text{shifted}}(x) = f(x+1)$.

If data are collected one-by-one in list mode, the shift can be followed as a function of time and an individual correction can be made for each count. The local shift is then obtained from an 'exponentially moving average':

$$\text{shift} \approx f(x) \left[- \int_a^b f(x) dx / \int_a^b f(x)^2 dx \right] n^{-1} + \text{shift}_{\text{previous}} (1 - n^{-1}) \quad (2)$$

in which n is a number representing the 'memory range'. A more stable trend line is obtained by analysing the data in the backward as well as the forward direction, as local effects cancel out in the average and global changes are detected with less delay.

3. Experimental result

Figure 1 shows a fraction of the summed ^{235}U alpha-particle spectrum, with and without gain shift correction. The software package SHIFTER (Version 1.1) automatically corrected for gain shifts between -5 and $+5$ channels, but needed some initial aid to correct for a shift by 20 channels when a new detector was taken into use. The best result was obtained with point-by-point shift correction.

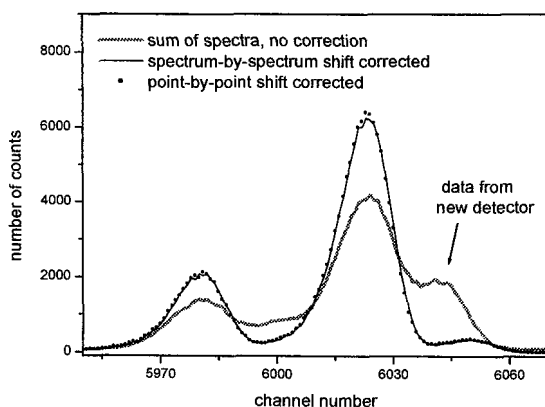


Figure 1. Gain shift corrected and non-corrected sum of 150 alpha spectra.

References

1. S. Pommé and G. Sibbens, *Appl. Radiat. Isot.*, in press.
2. D. O'Connor, G. Knoll and R. Ziff, *Nucl. Instr. and Meth.* **A353**, 291 (1994).

DEVELOPMENT OF SOFTWARE FOR ANOVA THAT CAN GENERATE EXPRESSIONS OF VARIANCE EXPECTATIONS

H. TANAKA AND K. EHARA

*National Institute of Advanced Industrial Science and Technology
AIST Tsukuba 3, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8563 Japan
E-mail: tanaka-hideyuki@aist.go.jp*

T. KAMOSHITA

*Ohken Associates
510 Syuwa TBR Bld., 2-10-2 Nagatacho, Chiyoda-ku, Tokyo 100-0014 Japan*

In recent years, a large number of programs have been equipped with ANOVA (Analysis of Variance) functions. The expression of expectation of variance in ANOVA must be calculated in order to evaluate each standard uncertainty. However, modern software does not yet offer the functionality to calculate this expression of expectation of variance. In this study expectations of variance in ANOVA were formulated and a new program that calculates the expression of the expectation of variance in typical and specific experimental designs and displays symbolic expectations of each variance was developed.

1. Introduction

There are two methods used in estimating the uncertainties of multiple factors. One focuses on a single factor, leaving levels for other factors fixed. This method is called a one-factor experiment. The other method uses combinations of every possible level for all factors. This method is called a full factorial experiment. The effects of a full factorial experiment are calculated by ANOVA, which is a tool often used in metrology. Annex H5 in GUM [1] is a first example of estimating uncertainty using ANOVA. ANOVA is used to estimate the effects of both within-day and between-day variability of observations. In recent years, a large number of programs have been equipped with ANOVA functions. However, ANOVA features of such software are designed mainly for F-tests and to determine the level of significance between each factor and error factor. Specifically, the expression of the expectation of variance of each factor must be calculated in order to calculate the variance of each factor. However, this kind of software is not yet equipped with a feature that allows calculation of the expression of expectation of variance.

This research focused on formulating expectations of variance in ANOVA [2] and developing a new program that calculates the expression of the expectation of variance in typical as well as specific experimental designs and displays symbolic expectations of each variance. This software greatly helps reduce work required in uncertainty evaluations.

2. Algorithms used in calculating expectation of variance

2.1. Completely randomized design

Completely randomized design is the method used when each piece of data measured is taken in random order. The following equation shows the expression of expectation of variance of factor B out of factors A, B, and C in a three-factor repetition experiment.

$$B: \sigma_e^2 + e_A e_C n \sigma_{A \times B \times C}^2 + e_A c n \sigma_{A \times B}^2 + e_C b n \sigma_{B \times C}^2 + a c n \sigma_B^2 \quad (1)$$

Here, e means that if factor A is a parameter, 0 is substituted for e_A and if factor A is a random variable, 1 is substituted for e_A . The level numbers are expressed by small letters of denotation and n is expressed by the number of repetitions. In Fig.1 we see an example of output made by our newly developed software. This is a three-factor repetition experiment in which factor A has 3 levels and is a parameter, factor B has 4 levels and is a random variable, and factor C has 5 levels and is also a random variable. Number of repetitions is 10.

	Expectation of Variance (Completely Randomized Design)
A	$\sigma_e^2 + 10 \sigma_{A \times B \times C}^2 + 40 \sigma_{A \times C}^2 + 50 \sigma_{A \times B}^2 + 200 \sigma_A^2$
B	$\sigma_e^2 + 30 \sigma_{A \times B \times C}^2 + 150 \sigma_B^2$
C	$\sigma_e^2 + 30 \sigma_{A \times B \times C}^2 + 120 \sigma_C^2$
AB	$\sigma_e^2 + 10 \sigma_{A \times B \times C}^2 + 50 \sigma_{A \times B}^2$
AC	$\sigma_e^2 + 10 \sigma_{A \times B \times C}^2 + 40 \sigma_{A \times C}^2$
BC	$\sigma_e^2 + 30 \sigma_{A \times B \times C}^2$
ABC	$\sigma_e^2 + 10 \sigma_{A \times B \times C}^2$
E4	σ_e^2

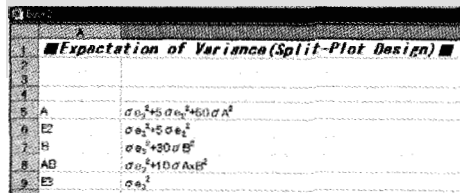
Fig.1: Example of a completely randomized Design

2.2. Split-plot design

In a split-plot design, measured data is not taken in a random order but randomization is performed hierarchically. The following equation shows the expression of expectation of variance of: firstly, factor A when factor A is (parameter, a levels, number of repetition: n times), and secondly B when it is (parameter, b levels) in a split-plot design.

$$A: \sigma_{e3}^2 + b \sigma_{e3}^2 + e_B n \sigma_{A \times B}^2 + b n \sigma_A^2 \quad (2)$$

Fig.2 shows an example of output by our software. This is the expectation of variance in factors A (parameter, 3 levels, $n=10$) and B (parameter, 5 levels).



The screenshot shows a software window titled "Expectation of Variance(Split-Plot Design)". It contains a table with two columns: factor labels and their corresponding variance expressions.

A	$\sigma e^2 + 5 \sigma e_1^2 + 50 \sigma A^2$
E2	$\sigma e_1^2 + 5 \sigma e_2^2$
B	$\sigma e_1^2 + 30 \sigma B^2$
AB	$\sigma e_1^2 + 10 \sigma A B^2$
E3	σe_2^2

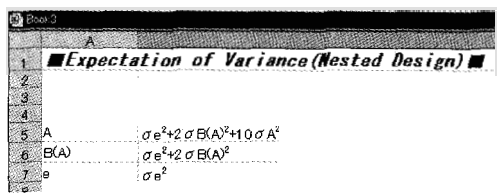
Fig.2: Example of split-plot design

2.3. Nested design

Nested design is one method used when factors are random variables. The following equation shows the expression of expectation of variance of factors A, B, and C in a nested design.

$$A : \sigma^2_{C(AB)} + c\sigma^2_{B(A)} + bc\sigma^2_A \qquad B : \sigma^2_{C(AB)} + c\sigma^2_{B(A)} \qquad C : \sigma^2_{C(AB)} \qquad (3)$$

Fig.3 is an example of output produced by our software. This is the expectation of variance in factors A (3 levels), B (5 levels), and C (2 levels).



The screenshot shows a software window titled "Expectation of Variance(Nested Design)". It contains a table with two columns: factor labels and their corresponding variance expressions.

A	$\sigma e^2 + 2 \sigma B(A)^2 + 10 \sigma A^2$
B(A)	$\sigma e^2 + 2 \sigma B(A)^2$
e	σe^2

Fig.3: Example of nested design

3. Conclusions

In this research, we produced software that automatically calculates not only each variance but the expression of the expectation of variance for a full factorial experiment as well as in split-plot and nested designs. This software works as an add-in for Microsoft Excel. In the future there are plans to enable free download of this software.

References

- 1. BIPM, IEC, IFCC, IUPAC, IUPAP, and OIML. Guide to the Expression of Uncertainty in Measurement, ISBN 92-67-10188-9, Second Edition (1995).
- 2. Y. Kondo, and W. Funasaka, Statistical Method for Engineers, (In Japanese), KYORITSU SHUPPAN CO., LTD., (1967).

TEMPLATE MATCHING IN PASSIVE SONAR RECOGNITION*

J. L. TERRY¹, D.A.TURNER² AND J. C. MASON¹

¹*School of Computing and Engineering, University of Huddersfield,
Huddersfield, UK. E-mail: j.l.terry@hud.ac.uk*

²*Thales Underwater Systems, Cheadle Heath, Stockport, UK.*

Passive sonar records the sound radiated by a target. A signal corresponding to a vessel is called a frequency track. A single noise source typically produces several frequency tracks, which collectively form a harmonic set. Often, there are several detectable harmonic sets, and so a key problem in passive sonar is to group the frequency tracks into their corresponding harmonic sets. This paper describes a novel method of identifying the different harmonic sets using a combination of data fitting and template matching.

1. Introduction

This paper describes an algorithm for associating into groups frequency tracks recorded using passive sonar. A frequency track is a noise signal corresponding to a target, typically there are several tracks associated to each noise source. Each group of frequency tracks is called a harmonic set, and a harmonic set provides vital information to a sonar operator in classifying a target. This paper uses the magnitude FFT data from the measured noise, which gives frequency and amplitude information over time. The frequency tracks in the harmonic set have the property that if the lowest frequency track (which is called the fundamental) has a frequency $f(t)$ at time t , then all other tracks have a frequency $\lambda f(t)$, where λ is a positive integer value.

The algorithm described in this paper uses a technique known as template matching [2] in order to construct the harmonic sets, but also exploits the known physical behaviour of frequency tracks belonging to a harmonic set. Typically, template matching is used only to compare how closely two objects match, but here it is also necessary to make a decision as to whether the frequency tracks should be associated.

* Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

2. Analysis Procedure

Figure 1 outlines the analysis procedure of each batch of data. Template matching [2] is used to provisionally group frequency tracks into harmonic sets. The first step is to form a template, which is achieved by fitting a B-spline $g(t)$ [1] to a frequency track. A simple knot placement algorithm that places knots at every m th data point is used, where the number of knots placed determines n .

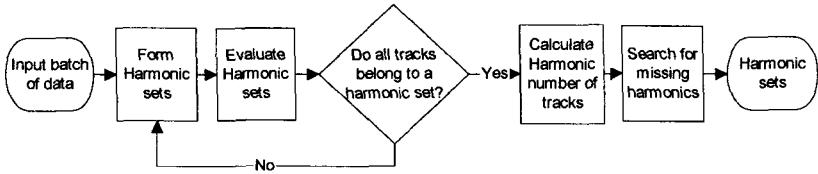


Figure 1 Overview of analysis procedure for each batch of data.

The physics governing harmonic sets mean that the only transformation required to successfully match tracks from the same set is a scaling in frequency, defined as S . Each data set, representing a frequency track can be expressed as

$$\{(f_i, t_i)\}_{i=1}^m = (f, t), \quad (1)$$

where f_i represents the frequency of the data at time t_i . Template matching transforms (f, t) to a new data set (\hat{f}, t) in the same region as $g(t)$, where $\hat{f} = Sf$.

The quality of match is measured in the l_2 -norm, so that the optimum match is found by solving the least squares minimisation problem

$$\min_S \left\{ E = (\hat{f} - g(t))^T (\hat{f} - g(t)) \right\}. \quad (2)$$

The value of S that achieves this minimisation is found by differentiating E with respect to S and setting the result equal to zero. This yields the result

$$f S = g(t), \quad (3)$$

which can be solved using standard linear least squares algorithms.

A decision is required as to whether a positive match has occurred, since tracks may be present that do not match the template. The decision is made using the value of the residual error E , if E is less than some user-defined threshold then the match is found to be positive, and the track is provisionally assigned to the harmonic set. If several harmonic sets exhibit similar characteristics then this technique can result in tracks being incorrectly assigned to a set. The physics governing harmonic sets mean that there is a relationship

between the scaling factors, S , calculated for the frequency tracks in the set that can be exploited to identify any rogue tracks. The scaling factor, S , between the template and the n th harmonic in a set is $S = n/N$, where N is the harmonic number of the template. The difference in scaling factors of sequential harmonics is constant and equals

$$S_{N \rightarrow n+1} - S_{N \rightarrow n} = (n+1)/N - n/N = 1/N. \quad (4)$$

Assuming there are more correctly assigned tracks than rogue tracks in the harmonic set, the value $1/N$, termed the difference in scaling factors δ , can be found by calculating the differences in scaling factors for every pairwise combination in the set. The most frequent value within a given tolerance will equal δ . All tracks with another track with a value δ between their scaling factors are identified as harmonics of that set, all other tracks are discarded.

3. Numerical Results

Figure 2a. shows a grayscale plot of a set of synthetic frequency data. At any given time, the brightness of the pixel image increases as the amplitude of the frequency increases. A frequency track is distinguished by high amplitude frequency data, and so from Figure 2a it can be seen that there are three different harmonic sets and corresponding frequency tracks. Figure 2b. shows the harmonic sets that have been formed for the data set. In the figure, frequency tracks belonging to different harmonic sets are distinguished by being plotted with different line styles. It can be seen from Figure 2b that the algorithm successfully constructs the harmonic sets.

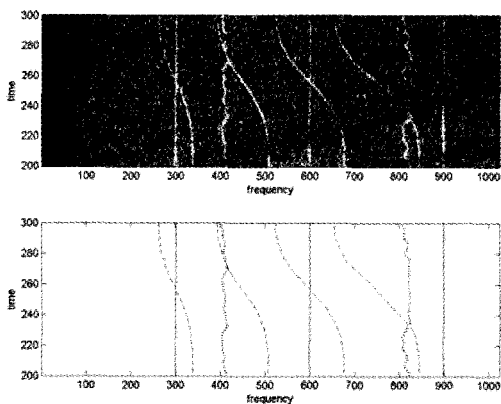


Figure 2 a.) Sample batch of data. b.) The resulting harmonic sets.

References

1. Dierckx, P. *Curve and Surface Fitting with Splines*. Oxford University Press, Oxford, UK, 1993.
2. Turner, D. A. *The Approximation of Cartesian Coordinate Data by Parametric Orthogonal Regression*. Ph.D. thesis, School of Computing and Mathematics, University of Huddersfield, Huddersfield, UK, 1999.

FAST COMPUTATIONAL ALTERNATIVE TO MONTE CARLO SIMULATION OF AN OUTPUT DISTRIBUTION

R. WILLINK

*Applied Mathematics, Industrial Research Ltd.,
P.O. Box 31-310, Lower Hutt, New Zealand
E-mail: r.willink@irl.cri.nz*

B. D. HALL

*Measurement Standards Laboratory of New Zealand,
Industrial Research Ltd., P.O. Box 31-310, Lower Hutt, New Zealand
E-mail: b.hall@irl.cri.nz*

We describe a method for the rapid calculation of an interval covering approximately a specified proportion of the distribution of a function of independent random variables. Its speed will make it particularly valuable in inverse problems. The first four moments of the output variable are obtained by combining the moments of the input variables according to the function. A Pearson distribution is fitted to these moments to give 95% intervals that are accurate in practice.

1. Introduction

Suppose that a known function F relates a measurand to n influence quantities. Let X_1, \dots, X_n be variables with the distributions assigned to the n input quantities and let Y represent the measurand, so $Y = F(X_1, \dots, X_n)$. This paper describes a rapid method of obtaining an interval enclosing approximately a specified proportion, say 0.95, of the mass of the distribution of Y . The method can serve as a complement to the method of Monte Carlo simulation, and may be particularly valuable in inverse problems where an interval is specified and a distribution is sought.

2. Moments and cumulants

The method involves calculation with the moments and cumulants of the variables. The r th moment about the origin of any variable X is $\mu'_r(X) \equiv E[X^r]$ and the r th central moment is $\mu_r(X) \equiv E[(X - \mu'_1)^r]$.

These are related by the equations¹

$$\begin{aligned}\mu_r &= \sum_{j=0}^r (-1)^j \binom{r}{j} (\mu'_1)^j \mu'_{r-j} \\ \mu'_r &= \sum_{j=0}^r \binom{r}{j} (\mu'_1)^j \mu_{r-j}.\end{aligned}$$

The mean μ'_1 and variance μ_2 determine location and scale, while μ_3 and μ_4 add information on skewness $\sqrt{\beta_1} \equiv \mu_3/\mu_2^{3/2}$ and kurtosis $\beta_2 \equiv \mu_4/\mu_2^2$.

Cumulants $\{\kappa_r\}$ are functions of moments, and vice versa.¹ They are related by²

$$\begin{aligned}\kappa_r &= \mu_r - \sum_{j=2}^{r-2} \binom{r-1}{j} \mu_j \kappa_{r-j} & r \geq 2 \\ \mu_r &= \kappa_r + \sum_{j=2}^{r-2} \binom{r-1}{j} \mu_j \kappa_{r-j} & r \geq 2.\end{aligned}$$

In particular, $\kappa_1 = \mu'_1$, $\kappa_2 = \mu_2$, $\kappa_3 = \mu_3$, $\kappa_4 = \mu_4 - 3\mu_2^2$.

3. Moments of input distributions

The symmetric input variables listed in Table 1 are the arc-sine variable on the interval from -1 to 1 , $A[-1, 1]$; the triangular variable, $T[-1, 1]$; the uniform variable, $U[-1, 1]$; the standard normal variable, $N(0, 1)$; and the Student's t -variable, t_ν . Exponential variables can also be used.

Table 1. Moments of symmetric input distributions.

Input variable	Moment μ_r (r even)	Kurtosis β_2
$A[-1, 1]$	$\frac{1 \cdot 3 \dots (r-1)}{2 \cdot 4 \dots r}$	1.5
$U[-1, 1]$	$1/(r+1)$	1.8
$T[-1, 1]$	$\frac{2}{(r+1)(r+2)}$	2.4
$N(0, 1)$	$\frac{r!}{2^{r/2} (r/2)!}$	3
t_ν	$\frac{\nu^{r/2} \cdot 1 \cdot 3 \dots (r-1)}{(\nu-2)(\nu-4) \dots (\nu-r)}$	$3 + \frac{6}{\nu-4}$

4. Finding the moments of the output

To obtain the moments of Y we break up F into elementary operations $+$, $-$, \times , \div , \exp , \log etc., and carry out moment calculations for each operation, while converting between moments and cumulants as required.

Arithmetic operations on independent variables X_h and X_i obey

$$\begin{aligned}\kappa_r(aX_h + bX_i) &= a^r \kappa_r(X_h) + b^r \kappa_r(X_i) \\ \mu'_r(aX_h X_i) &= a^r \mu'_r(X_h) \mu'_r(X_i).\end{aligned}$$

Moments about zero quickly become large if the mean is distant from zero and, consequently, round-off errors dominate in sums of large oscillating terms. So the higher moments of a product are actually obtained by expressing $\mu_r(aX_hX_i)$ in terms of the means and central moments,

$$\mu_r(aX_hX_i) = a^r \sum_{k=0}^r \sum_{j=0}^k \frac{r! \mu'_1(X_h)^{k-j} \mu'_1(X_i)^j \mu_{r-k+j}(X_h) \mu_{r-j}(X_i)}{(r-k)!(k-j)!j!},$$

which is obtained by writing X_h and X_i each as its mean plus a deviation.

Let μ'_1 and μ_r denote $\mu'_1(X)$ and $\mu_r(X)$. For the reciprocal and other functions $f(X)$ we have the Taylor expansion $f(X) = f(\mu'_1) + \sum_{j=1}^{\infty} c_j \delta^j$ where $c_j \equiv f^{(j)}(\mu'_1)/j!$ and $\delta = X - \mu'_1$. Analysis gives

$$\begin{aligned} \mu'_1(f(X)) &= f(\mu'_1) - \phi & \phi &\equiv -\sum_{j=2}^{\infty} c_j \mu_j \\ \mu_r(f(X)) &= r! \sum_{q_0=0}^r \phi^{q_0} L(r-q_0)/q_0! \end{aligned}$$

where

$$L(r-q_0) \equiv \sum_{q_1=0}^{r-q_0} \sum_{q_2=0}^{r-q_0-q_1} \sum_{q_3=0}^{r-q_0-q_1-q_2} \cdots \left(\prod_{j=1}^{\infty} \frac{c_j^{q_j}}{q_j!} \right) \mu_Q \quad \sum_{j=0}^{\infty} q_j = r,$$

with $Q \equiv \sum_{j=1}^{\infty} j q_j$. The c_j values are easily found for the common functions $f(X) = 1/X, \sqrt{X}, 1/\sqrt{X}, X^b, \log X, \exp X, \cos X, \sin X$. The greatest practical difficulty in the calculation of $\mu_r(f(X))$ is in the calculation of $L(r-q_0)$, which, in the general case, is a convergent sum of infinite dimension. An algorithm for this has been obtained.

If X is uniformly distributed between the positive limits a and b then exact results are

$$\begin{aligned} \mu'_r(\ln X) &= \sum_{k=0}^r (-1)^k \frac{r!}{(r-k)!} \cdot \frac{b \ln^{r-k} b - a \ln^{r-k} a}{b-a} & (\text{Ref. 3, Eq. 2.722}) \\ \mu'_r(e^X) &= \frac{e^{rb} - e^{ra}}{r(b-a)} \\ \mu'_r(\sqrt{X}) &= \frac{2}{r+2} \cdot \frac{b^{r/2+1} - a^{r/2+1}}{b-a} \\ \mu'_r(1/X) &= \frac{1}{b-a} \cdot \frac{a^{1-r} - b^{1-r}}{r-1} & r \geq 2 \\ \mu'_1(1/X) &= \frac{\ln b - \ln a}{b-a}. \end{aligned}$$

Our implementation of the method involves an algorithm that keeps track of the moments of the X_i variables and variables formed at intermediate stages in the calculation of Y .

5. Approximating with a Pearson distribution

The Pearson system of statistical distributions includes normal, uniform, Student's t , chi-square, gamma, beta and arc-sine distributions as special cases.⁴ The system contains a unique distribution for every possible set of the moments μ'_1 , μ_2 , μ_3 and μ_4 , for matching to the same moments of Y . The flexibility of the system means that the distribution obtained will be a good approximation to the distribution of Y in practice.

We find the interval enclosing the specified proportion of this Pearson distribution from its standardized form. Percentage points of standardized Pearson distributions indexed by β_1 (or $\sqrt{\beta_1}$) and β_2 are available in tables for certain probability levels,^{5,6} or may be calculated by a simple approximate method for an arbitrary probability level.^{7,8} For greater accuracy, the percentage points could be found from the distribution itself, whose probability density function can be obtained from the moments.⁴

6. Examples

(i) Suppose $Y = \ln X_1 \times \exp X_2$ where $X_1 \sim U[9, 11]$ and $X_2 \sim U[3, 3.1]$ independently. (The notation \sim means 'is distributed as'.) The method gives $\mu'_1(Y) = 48.569$, $\mu_2(Y) = 3.476$, $\sqrt{\beta_1}(Y) = 0.0314$ and $\beta_2(Y) = 2.389$. By interpolating in the tables,⁵ we find that the interval containing the central 95% of the Pearson distribution with this set of moments is $[45.1, 52.1]$. Monte Carlo simulation using 10^6 trials gave the interval $[45.123 \pm 0.007, 52.188 \pm 0.007]$, where the uncertainties represent 2 standard deviations.

(ii) Suppose $Y = X_1 X_2 + X_3$ where $X_1 \sim N(10, 4)$, $X_2 \sim N(10, 1)$ and $X_3 \sim U[90, 110]$ independently. We find $\mu'_1(Y) = 200$, $\mu_2(Y) = 537.33$, $\sqrt{\beta_1}(Y) = 0.193$ and $\beta_2(Y) = 3.08$, and the 95% interval of Y is $[157, 248]$.⁵ Monte Carlo simulation with 10^6 trials gave $[156.6 \pm 0.1, 247.5 \pm 0.2]$.

7. Comparison of speed of computation

The advantage in speed over Monte Carlo simulation for a similar level of accuracy depends on the function and inputs, and is not trivial to assess.

For example, let $Y = X_1 X_2$ where X_1 and X_2 are independent normal variables. In this situation, a 95% interval is obtained in the same time as that required by a Monte Carlo simulation involving only 100 *unsorted* trials. If $X_1 \sim N(3, 1)$ and $X_2 \sim N(2, 1)$ then the calculated interval covers 0.953 of the true distribution of Y , and only 30% of Monte Carlo simulations involving 10^3 trials give intervals that cover a proportion closer

to 0.95. However, if $X_1 \sim N(10, 4)$ and $X_2 \sim N(10, 1)$ then the calculated interval covers 0.9501 of the true distribution of Y , while only 40% of Monte Carlo intervals involving 10^6 trials cover a proportion that is closer to 0.95. So, on average, Monte Carlo required much more time to achieve the same accuracy, especially in the second case. Figure 1 depicts these two cases.

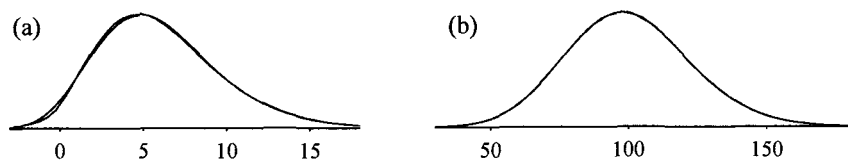


Figure 1. True and approximate distributions of X_1X_2 ; (a) $X_1 \sim N(3, 1)$, $X_2 \sim N(2, 1)$; (b) $X_1 \sim N(10, 4)$, $X_2 \sim N(10, 1)$. In (a) the distributions can be distinguished between $x = 0$ and $x = 5$. In (b) they are indistinguishable.

8. Comments

The method requires the input variables to be independent, so correlated variables have to be treated to put the problem in the required form. For example, if X_1 and X_2 in example (ii) are dependent then we write $Y = X_4 + X_3$ where $X_4 \equiv X_1X_2$, and the technique can be applied if the moments of X_4 are calculable by some other means.

The technique is straightforward in theory but is challenging to implement. For example, an aspect not discussed here is the slight truncation of variables with infinite tails in order to avoid numerical problems that are caused by large or undefined high-order central moments.

The method is the subject of patent application PCT/NZ02/00228.

References

1. A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, Vol.1, 5th ed., Griffin, (1987).
2. R. Willink, *Comm. Statist. Theory Methods* **32**, 701 (2003).
3. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* 6th ed. Academic Press. (2000).
4. W. P. Elderton and N. L. Johnson, *Systems of Frequency Curves*, Cambridge University Press, (1969).
5. E. S. Pearson and H. O. Hartley eds., *Biometrika Tables for Statisticians*, Vol. 1, 3rd ed. Cambridge University Press, (1966). [Table 42].
6. E. S. Pearson and H. O. Hartley eds., *Biometrika Tables for Statisticians*, Vol. 2. Cambridge University Press, (1972). [Table 32].
7. K. O. Bowman and L. R. Shenton, *Biometrika* **66**, 147 (1979).
8. C. S. Davis and M. A. Stephens, *J. Roy. Statist. Soc. Ser. C* **32**, 322 (1983).

SHORT COURSE ON UNCERTAINTY EVALUATION*

M. G. COX

*Centre for Mathematics and Scientific Computing
National Physical Laboratory, Queens Road, Teddington, UK
Email: maurice.cox@npl.co.uk*

A short course on uncertainty evaluation was held in association with the international conference on Advanced Mathematical and Computational Tools for Metrology (AMCTM 2003). The objectives of the course are stated and a brief description given of the lessons learned.

1. Objectives

The objectives of the course were to

- Provide the probabilistic basis of uncertainty evaluation
- Present a formulation of the generic problem of uncertainty evaluation that is consistent with this basis
- Provide implementations (computational tools) for solving this problem based on the law of propagation of uncertainty (LPU) and the propagation of distributions
- Apply these implementations to uncertainty evaluation problems arising in several branches of metrology, and compare them
- Present and show the role of computational tools such as bootstrap re-sampling in the context of uncertainty evaluation
- Emphasize the alignment of the coverage to the Guide to the Expression of Uncertainty in Measurement (GUM)¹ and the supplemental guides being developed to support it.

2. Content

In order to fulfill these objectives, the content of the course was

*Work partially funded under EU project SofTools.MetroNet Contract N. G6RT-CT-2001-05061 and supported by the Software Support for Metrology programme of the UK's Department of Trade and Industry.

- (1) The probabilistic basis of uncertainty evaluation in metrology. Information-based probability and probability density function (PDF). Model of measurement. Propagation of PDFs. Coverage interval and coverage probability. Estimates of measurands and associated uncertainty matrix. Assigning a PDF to a measurand.
- (2) Implementation of uncertainty evaluation I. LPU. Multiple (output) measurands. Implicit models. Mutually dependent input quantities. Use of matrices.
- (3) Implementation of uncertainty evaluation II. The propagation of distributions. Realization using Monte Carlo simulation (MCS). Interpretation of results.
- (4) Applications to various metrology areas. LPU. Propagation of distributions. Comparisons.
- (5) Computational tools. Type A uncertainty evaluation. Bootstrap re-sampling. Standard uncertainties. Coverage intervals.

3. Course tutors

The tutors were Dr Patrizia Ciarlini (IAC-CNR, Rome), Professor Maurice Cox (NPL), Dr Peter Harris (NPL), Dr Bernd Siebert (PTB) and Dr Wolfgang Wöger (PTB), experts in various areas of uncertainty evaluation.

4. Lessons learned

Lessons learned from the course were

- (1) The GUM is a rich document, but the recommendations contained within it are not always applicable, a statement applying both to users and developers of software for uncertainty evaluation. Software providers should be aware that some practitioners may make use of software that may not be fit for all purposes to which it is applied. Tests for the applicability of the approaches implemented in the software would be desirable for quality assurance reasons.
- (2) The GUM contains anomalies, but mainly provides a consistent and logical probabilistic basis for uncertainty evaluation. There are advantages in applying this probabilistic approach to all uncertainty evaluation problems, regardless of the extent of their complexity.
- (3) The basis provides a summary of the knowledge of the values of the input and output quantities in a model of measurement in the form of probability density functions (PDFs) or joint PDFs.

- (4) Within the probabilistic approach uncertainties are not estimated *per se*, but are calculated from the PDFs for the values of the relevant quantities.
- (5) Bayes' Theorem and the principle of maximum (information) entropy are valuable in the assignment of PDFs on the basis of available information concerning the values of the input quantities.
- (6) Two distinct types of correlation occur in practice: statistical and logical. The former arises when measurements are taken pair-wise, and the latter when commonality exists in the form of measurements taken, e.g., with the same instrument.
- (7) A coverage interval, corresponding to a prescribed coverage probability, for the value of the output quantity always exists for a given PDF, even though moments of that PDF (e.g., the expectation as a best estimate of the value of the output quantity, and the standard deviation as the associated standard uncertainty) may not always exist. The Cauchy distribution is an example of such a PDF.
- (8) There would be advantage in avoiding the concept of effective degrees of freedom in any new edition of the GUM, since it is inconsistent with the probabilistic basis. It was noted that, according to the Joint Committee for Guides in Metrology (JCGM)^a, which has responsibility for the GUM, the GUM will not be revised *per se* in the foreseeable future. Rather, supplemental guides to the GUM are being prepared by the JCGM to address such issues.
- (9) The use of the propagation of distributions (as opposed to LPU) provides, in principle, the PDF for the value of the output quantity. MCS furnishes a straightforward implementation of the propagation of distributions. Such an implementation can be used in any specific instance to validate the use of LPU, and the application of the Central Limit Theorem as a basis for establishing a coverage interval for the value of the output quantity.

References

1. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML. Guide to the Expression of Uncertainty in Measurement, 1995. ISBN 92-67-10188-9, Second Edition.

^a<http://www.bipm.org/en/committees/jc/jcgm/>

SOFTWARE REQUIREMENTS IN LEGAL METROLOGY SHORT COURSE HELD ADJACENT TO THE CONFERENCE^{*c}

D. RICHTER

*Physikalisch-Technische Bundesanstalt Braunschweig and Berlin
Abbestr. 2-12, 10587 Berlin, Germany
E-mail: dieter.richter@ptb.de*

In recent years, a large number of programs have been equipped with ANOVA (Analysis of Variance) functions. The expression of expectation of variance in ANOVA must be calculated in order to evaluate each standard uncertainty. However, modern software does not yet offer the functionality to calculate this expression of expectation of variance. In this study expectations of variance in ANOVA were formulated and a new program that calculates the expression of the expectation of variance in typical and specific experimental designs and displays symbolic expectations of each variance was developed.

1. Aim of the course

Results of the European Network “MID-Software”^d, in which a framework and sets of particular software requirements have been developed, were presented. The requirements are harmonized among the participating European partners from Notified Bodies and industry. Although they are based on the European “Measuring Instruments Directive” (MID) and related to applications in legal metrology, there are principles and even particular requirements, which can be transferred to non-legal environments. The course was aimed at presenting the approach and explaining the particular sets of requirements developed.

2. Requirements and validation

Considering the current discussion on software validation in metrology, the definition of requirements plays an essential role. Validation means provision of

^{*} Work partially funded under EU SofTools_NetroNet Contract N° G6RT-CT-2001-05061.

^c In addition to the author, the course was given by Daryoush Talebi, Ulrich Grottker (both PTB, Germany) and Mike Fortune (NWML, UK).

^d The work is funded by the European commission under the registration number G7RT-CT-2001-05064

evidence that particular requirements are fulfilled. This means that requirements must be clear when a validation is going to be carried out. The importance of harmonization of requirements is underlined by the fact that similar requirements can be verified by similar methods. Insofar, the work undertaken in legal metrology is a good example of how software validation can be organized on the basis of task division.

The focus of the software requirements in legal metrology is the security of software and data, i.e. the protection of software and measurement data from accidental or intended changes. This is an issue of consumer protection, which is a basic objective of legal metrology. However, software security is an issue of legal metrology not only but concerns software in metrology in general. It represents one class of requirements of metrology software. Further classes are

- software requirements with respect to conformity of underlying models,
- software requirements with respect to conformity of standards, etc.,
- software requirements with respect to functional correctness
- software requirements with respect to quality of implementation,
- software requirements with respect to numerical stability,
- software requirements with respect to data acquisition,
- software requirements with respect to data transmission,
- software requirements with respect to performance characteristics,
- software requirements with respect to ergonomic criteria.

3. Objective of the European Network “MID-Software”

The Network aims at supporting the implementation of the European Measuring Instrument Directive (MID) in the software area by removing uncertainty as to the interpretation of software requirements and by building mutual confidence in the results of software testing. The main objective is the harmonization of MID implementations with respect to software in all EU member states.

The work is proposed to lead to harmonized software requirements laid down as guidance for manufacturers and, as the requirements are simultaneously the “technical interface” between manufacturers and Notified Bodies, for comparable assessments by Notified Bodies in the various European member states

4. Definition of requirements

The original objective was to define and differentiate the software requirements according to different categories of measuring instruments. At an early stage of the work, it seemed to be more advantageous to develop the requirements on a rather generic basis. The basis is given by typical device-independent configurations of measuring instruments instead of instrument categories. Furthermore, device-independent risk classes are introduced to which measuring instruments are assigned.

As regards the configuration, the requirements are assigned to two main groups:

- generic software requirements of basic instruments,
- generic software requirements of extended IT functionality.

Two classes are differentiated as regards software requirements of basic instruments:

- software of built-for-purpose instruments,
- software of instruments that use a universal computer.

To show the different conditions the definition, typical characteristics of the two classes are given.

A built-for-purpose instrument is a measuring instrument with an embedded software system. It is characterized by the following features:

- The entire software of the measuring instrument is constructed for the measuring purpose.
- The software normally is not intended to be modified after approval.
- The user interface is dedicated to the measuring purpose.
- The software is invariable and there are no means for programming or loading software.
- The device may have interfaces for data transmission. The respective requirements have to be observed.
- The device may have the ability to store measurement data. The respective requirements have to be observed.

Requirements for software of instruments that use a universal computer must observe the following features in particular:

- Any operating system and supporting software (e.g. drivers) may be used.

- In addition to the measuring instrument application, other software applications may at the same time also reside on the system.
- The computer system may standalone, be part of a closed network, or be connected to open networks.
- Storage of measurement data may be local or remote, fixed or removable.
- The user interface may be switched from an operating mode dedicated to a measuring function to a mode, which is not, and vice versa.
- As the system is general-purpose, the measuring sensor would normally be external to the computer module and be linked to it by a communication link.

The second main group of requirements relates to the extended IT functionality of measuring instruments. There are four different groups of requirements:

- software requirements for the storage of measurement data,
- software requirements for the transmission of measurement data,
- software requirements for the software download,
- software requirements for the separation of software into legally relevant and legally non-relevant data.

The requirements for these groups apply only if the device under consideration has a respective function.

References

The work reported on is still in progress. The most recent developments, drafts and news are available under www.mid-software.org.

Author Index

- Allasia, G. 137
Antunes, S. D. 1
Aranda, S. 149
- Bachmann, J. 149
Bich, W. 159, 274
Boudjemaa, R. 67, 170
Bourdet, P. 82
Bremsen, W. 16
- Callegaro, L. 274
Castelletto, S. 279
Chunovkina, A. 180
Ciarlini, P. 24, 122, 293
Codda, M. 106
Cox, M. G. 37, 52, 67, 170, 342
Crampton, A. 67
Crenna, F. 106
- D'Errico, G. E. 282
Degiovanni, I. P. 279
Dematteis, R. 323
Douglas, R. J. 264
- Ehara, K. 330
- Filipe, E. 285
Flowers, J. L. 37
Forbes, A. B. 37, 67, 170
Fotowicz, P. 302
- Haitjema, H. 299
Hald, J. 189
Hall, B. D. 199, 337
Harris, P. M. 37, 170
Hässelbarth, W. 16
Hetman, A. 302, 309
Hill, K. D. 264
- Ichim, D. 209, 293
Iuculano, G. 219
- Jenkinson, D. P. 67
- Kamoshita, T. 330
Kamvissis, S. 52
Korczynski, M. J. 302, 309
- Lartigue, C. 82
Lei, D. 314
Linares, J. M. 149
Lo Cascio, M. L. 24
- Marcarino, P. 318, 323
Mason, J. C. 67, 314, 333
Mathieu, L. 82
Mehdi-Souzani, C. 82
Merlone, A. 318, 323
Milton, M. J. T. 52
Morel, M. A. A. 299
- Nielsen, L. 189
- Panfilio, G. 229
Pavese, F. 240, 293
Pellegrini, G. 219
Penneccchi, F. 159, 274
Peroni, I. 209
Perruchet, C. 98
Pommé, S. 327
Premoli, A. 251, 293
- Rastello, M. L. 251, 279, 293
Richter, D. 345
Rossi, G. B. 106
Ruo Berchera, I. 279
- Sibbens, G. 327
Siebert, B. R. L. 122
Sparasci, F. 209
Sprauel, J. M. 149
Steele, A. G. 264
Steuer, P. P. M. 318, 323
Szmyrka-Grzebyk, A. 302

Tanaka, H. 330

Tavella, P. 229

Terry, J. L. 333

Turner, D. A. 333

Vargha, G. 52

Vicente, M. A. F. 1

Willink, R. D. 337

Zanobini, A. 219

Zucca, C. 229

Series on Advances in Mathematics for Applied Sciences

Editorial Board

N. Bellomo

Editor-in-Charge

Department of Mathematics
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino
Italy
E-mail: nicola.bellomo@polito.it

F. Brezzi

Editor-in-Charge

Istituto di Analisi Numerica del CNR
Via Abbiategrasso 209
I-27100 Pavia
Italy
E-mail: brezzi@imati.cnr.it

M. A. J. Chaplain

Department of Mathematics
University of Dundee
Dundee DD1 4HN
Scotland

C. M. Dafermos

Lefschetz Center for Dynamical Systems
Brown University
Providence, RI 02912
USA

J. Felcman

Department of Numerical Mathematics
Faculty of Mathematics and Physics
Charles University in Prague
Sokolovska 83
18675 Praha 8
The Czech Republic

M. A. Herrero

Departamento de Matematica Aplicada
Facultad de Matemáticas
Universidad Complutense
Ciudad Universitaria s/n
28040 Madrid
Spain

S. Kawashima

Department of Applied Sciences
Engineering Faculty
Kyushu University 36
Fukuoka 812
Japan

M. Lachowicz

Department of Mathematics
University of Warsaw
Ul. Banacha 2
PL-02097 Warsaw
Poland

S. Lenhart

Mathematics Department
University of Tennessee
Knoxville, TN 37996-1300
USA

P. L. Lions

University Paris XI-Dauphine
Place du Marechal de Lattre de Tassigny
Paris Cedex 16
France

B. Perthame

Laboratoire d'Analyse Numerique
University Paris VI
tour 55-65, 5ieme etage
4, place Jussieu
75252 Paris Cedex 5
France

K. R. Rajagopal

Department of Mechanical Engrg.
Texas A&M University
College Station, TX 77843-3123
USA

R. Russo

Dipartimento di Matematica
Università degli Studi Napoli II
81100 Caserta
Italy

Series on Advances in Mathematics for Applied Sciences

Aims and Scope

This Series reports on new developments in mathematical research relating to methods, qualitative and numerical analysis, mathematical modeling in the applied and the technological sciences. Contributions related to constitutive theories, fluid dynamics, kinetic and transport theories, solid mechanics, system theory and mathematical methods for the applications are welcomed.

This Series includes books, lecture notes, proceedings, collections of research papers. Monograph collections on specialized topics of current interest are particularly encouraged. Both the proceedings and monograph collections will generally be edited by a Guest editor.

High quality, novelty of the content and potential for the applications to modern problems in applied science will be the guidelines for the selection of the content of this series.

Instructions for Authors

Submission of proposals should be addressed to the editors-in-charge or to any member of the editorial board. In the latter, the authors should also notify the proposal to one of the editors-in-charge. Acceptance of books and lecture notes will generally be based on the description of the general content and scope of the book or lecture notes as well as on sample of the parts judged to be more significantly by the authors.

Acceptance of proceedings will be based on relevance of the topics and of the lecturers contributing to the volume.

Acceptance of monograph collections will be based on relevance of the subject and of the authors contributing to the volume.

Authors are urged, in order to avoid re-typing, not to begin the final preparation of the text until they received the publisher's guidelines. They will receive from World Scientific the instructions for preparing camera-ready manuscript.

SERIES ON ADVANCES IN MATHEMATICS FOR APPLIED SCIENCES

- Vol. 17 The Fokker–Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solution
by C. Soize
- Vol. 18 Calculus of Variation, Homogenization and Continuum Mechanics
eds. G. Bouchitté et al.
- Vol. 19 A Concise Guide to Semigroups and Evolution Equations
by A. Belleni-Morante
- Vol. 20 Global Controllability and Stabilization of Nonlinear Systems
by S. Nikitin
- Vol. 21 High Accuracy Non-Centered Compact Difference Schemes for Fluid Dynamics Applications
by A. I. Tolstykh
- Vol. 22 Advances in Kinetic Theory and Computing: Selected Papers
ed. B. Perthame
- Vol. 23 Waves and Stability in Continuous Media
eds. S. Rionero and T. Ruggeri
- Vol. 24 Impulsive Differential Equations with a Small Parameter
by D. Bainov and V. Covachev
- Vol. 25 Mathematical Models and Methods of Localized Interaction Theory
by A. I. Bunimovich and A. V. Dubinskii
- Vol. 26 Recent Advances in Elasticity, Viscoelasticity and Inelasticity
ed. K. R. Rajagopal
- Vol. 27 Nonstandard Methods for Stochastic Fluid Mechanics
by M. Capinski and N. J. Cutland
- Vol. 28 Impulsive Differential Equations: Asymptotic Properties of the Solutions
by D. Bainov and P. Simeonov
- Vol. 29 The Method of Maximum Entropy
by H. Gzyl
- Vol. 30 Lectures on Probability and Second Order Random Fields
by D. B. Hernández
- Vol. 31 Parallel and Distributed Signal and Image Integration Problems
eds. R. N. Madan et al.
- Vol. 32 On the Way to Understanding The Time Phenomenon: The Constructions of Time in Natural Science — Part 1. Interdisciplinary Time Studies
ed. A. P. Levich
- Vol. 33 Lecture Notes on the Mathematical Theory of the Boltzmann Equation
ed. N. Bellomo

SERIES ON ADVANCES IN MATHEMATICS FOR APPLIED SCIENCES

- Vol. 34 Singularly Perturbed Evolution Equations with Applications to Kinetic Theory
by J. R. Mika and J. Banasiak
- Vol. 35 Mechanics of Mixtures
by K. R. Rajagopal and L. Tao
- Vol. 36 Dynamical Mechanical Systems Under Random Impulses
by R. Iwankiewicz
- Vol. 37 Oscillations in Planar Dynamic Systems
by R. E. Mickens
- Vol. 38 Mathematical Problems in Elasticity
ed. R. Russo
- Vol. 39 On the Way to Understanding the Time Phenomenon: The Constructions of Time in Natural Science — Part 2. The “Active” Properties of Time According to N. A. Kozyrev
ed. A. P. Levich
- Vol. 40 Advanced Mathematical Tools in Metrology II
eds. P. Ciarlini et al.
- Vol. 41 Mathematical Methods in Electromagnetism Linear Theory and Applications
by M. Cessenat
- Vol. 42 Constrained Control Problems of Discrete Processes
by V. N. Phat
- Vol. 43 Motor Vehicle Dynamics: Modeling and Simulation
by G. Genta
- Vol. 44 Microscopic Theory of Condensation in Gases and Plasma
by A. L. Itkin and E. G. Kolesnichenko
- Vol. 45 Advanced Mathematical Tools in Metrology III
eds. P. Ciarlini et al.
- Vol. 46 Mathematical Topics in Neutron Transport Theory — New Aspects
by M. Mokhtar-Kharroubi
- Vol. 47 Theory of the Navier–Stokes Equations
eds. J. G. Heywood et al.
- Vol. 48 Advances in Nonlinear Partial Differential Equations and Stochastics
eds. S. Kawashima and T. Yanagisawa
- Vol. 49 Propagation and Reflection of Shock Waves
by F. V. Shugaev and L. S. Shtemenko
- Vol. 50 Homogenization
eds. V. Berdichevsky, V. Jikov and G. Papanicolaou

SERIES ON ADVANCES IN MATHEMATICS FOR APPLIED SCIENCES

- Vol. 51 Lecture Notes on the Mathematical Theory of Generalized Boltzmann Models
by N. Bellomo and M. Lo Schiavo
- Vol. 52 Plates, Laminates and Shells: Asymptotic Analysis and Homogenization
by T. Lewiński and J. J. Telega
- Vol. 53 Advanced Mathematical and Computational Tools in Metrology IV
eds. P. Ciarlini et al.
- Vol. 54 Differential Models and Neutral Systems for Controlling the Wealth of Nations
by E. N. Chukwu
- Vol. 55 Mesomechanical Constitutive Modeling
by V. Kafka
- Vol. 56 High-Dimensional Nonlinear Diffusion Stochastic Processes — Modelling for Engineering Applications
by Y. Mamontov and M. Willander
- Vol. 57 Advanced Mathematical and Computational Tools in Metrology V
eds. P. Ciarlini et al.
- Vol. 58 Mechanical and Thermodynamical Modeling of Fluid Interfaces
by R. Gatignol and R. Prud'homme
- Vol. 59 Numerical Methods for Viscosity Solutions and Applications
eds. M. Falcone and Ch. Makridakis
- Vol. 60 Stability and Time-Optimal Control of Hereditary Systems — With Application to the Economic Dynamics of the US (2nd Edition)
by E. N. Chukwu
- Vol. 61 Evolution Equations and Approximations
by K. Ito and F. Kappel
- Vol. 62 Mathematical Models and Methods for Smart Materials
eds. M. Fabrizio, B. Lazzari and A. Morro
- Vol. 63 Lecture Notes on the Discretization of the Boltzmann Equation
eds. N. Bellomo and R. Gatignol
- Vol. 64 Generalized Kinetic Models in Applied Sciences — Lecture Notes on Mathematical Problems
by L. Arlotti et al.
- Vol. 65 Mathematical Methods for the Natural and Engineering Sciences
by R. E. Mickens
- Vol. 66 Advanced Mathematical and Computational Tools in Metrology VI
eds. P. Ciarlini et al.

SERIES ON ADVANCES IN MATHEMATICS FOR APPLIED SCIENCES

Published:

- Vol. 1 Mathematical Topics in Nonlinear Kinetic Theory II:
 The Enskog Equation
 by N. Bellomo et al.
- Vol. 2 Discrete Models of Fluid Dynamics
 ed. A. S. Alves
- Vol. 3 Fluid Dynamic Applications of the Discrete Boltzmann Equation
 by R. Monaco and L. Preziosi
- Vol. 4 Waves and Stability in Continuous Media
 ed. S. Rionero
- Vol. 5 A Theory of Latticed Plates and Shells
 by G. I. Pshenichnov
- Vol. 6 Recent Advances in Combustion Modelling
 ed. B. Larroturou
- Vol. 7 Linear Kinetic Theory and Particle Transport in Stochastic Mixtures
 by G. C. Pomraning
- Vol. 8 Local Stabilizability of Nonlinear Control Systems
 by A. Bacciotti
- Vol. 9 Nonlinear Kinetic Theory and Mathematical Aspects of
 Hyperbolic Systems
 ed. V. C. Boffi
- Vol. 10 Nonlinear Evolution Equations. Kinetic Approach
 by N. B. Maslova
- Vol. 11 Mathematical Problems Relating to the Navier-Stokes Equation
 ed. G. P. Galdi
- Vol. 12 Thermodynamics and Kinetic Theory
 eds. W. Kosiński et al.
- Vol. 13 Thermomechanics of Phase Transitions in Classical Field Theory
 by A. Romano
- Vol. 14 Applications of Padé Approximation Theory in Fluid Dynamics
 by A. Pozzi
- Vol. 15 Advances in Mathematical Modelling of Composite Materials
 ed. K. Z. Markov
- Vol. 16 Advanced Mathematical Tools in Metrology
 eds. P. Ciarlina et al.

